NAVAL TRAINING SYSTEMS CENTER
ORLANDO, FLORIDA

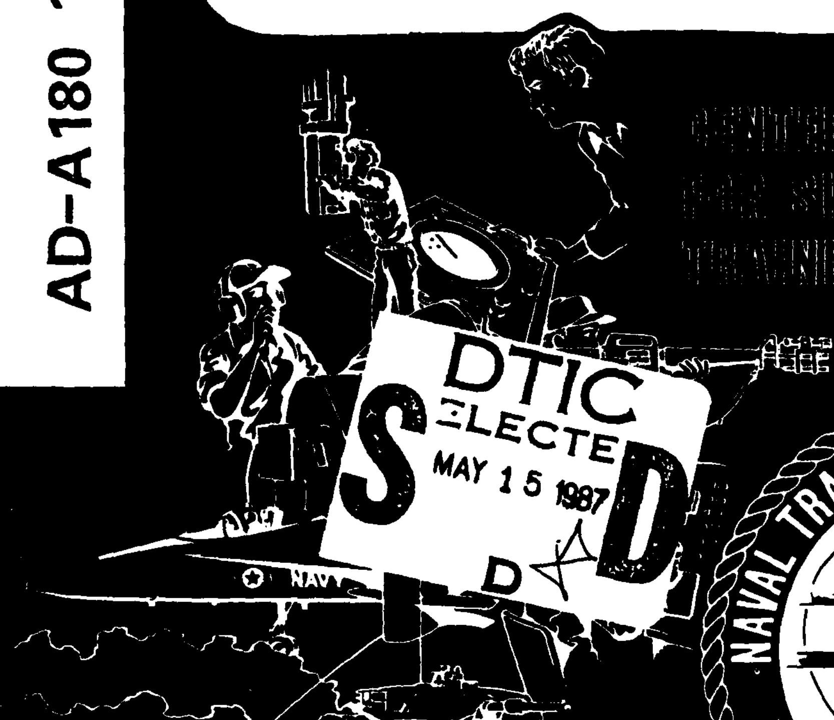
(12)

THE EFFECTS OF ASYNCHRONOUS VISUAL DELAYS ON SIMULATOR FLIGHT PERFORMANCE AND THE DEVELOPMENT OF SIMULATOR SICKNESS SYMPTOMATOLOGY

K. C. Uliano, E. Y. Lambert,
R. S. Kennedy, and D. J. Sheppard

Interim Final Report for the Period
April 25, 1986 - December 26, 1986
Contract No. N61339-85-D-0026

CENTER OF EXCELLENCE
FOR SIMULATION AND
TRAINING TECHNOLOGY



NAVAL TRAINING SYSTEMS CENTER
ORLANDO, FLORIDA

NAVTRASYSCEN 85-D-0026-1

(12)

THE EFFECTS OF ASYNCHRONOUS VISUAL DELAYS ON SIMULATOR FLIGHT PERFORMANCE AND THE DEVELOPMENT OF SIMULATOR SICKNESS SYMPTOMATOLOGY

K. C. Uliano, E. Y. Lambert,
R. S. Kennedy, and D. J. Sheppard

Interim Final Report for the Period
April 25, 1986 - December 26, 1986
Contract No. N61339-85-D-0026

Prepared by:
Essex Corporation
1040 Woodcock Road, Suite 227
Orlando, Florida 32803

26 December 1986

Prepared for:
Naval Training Systems Center
Orlando, Florida 32813-7100

NAVTRASYSCEN 85-D-0026-1

UNCLASSIFIED

AD-A180 196

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
NAVTRASYSCEN 86-D-0026-1

6a. NAME OF PERFORMING ORGANIZATION
Essex Corporation

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Naval Training Systems Center

6c. ADDRESS *(City, State, and ZIP Code)*
1040 Woodcock Road, Suite 227
Orlando, FL 32803

7b. ADDRESS *(City, State, and ZIP Code)*
Orlando, FL 32813-7100

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Naval Air Systems Command

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N61339-86-D-0026

8c. ADDRESS *(City, State, and ZIP Code)*
Washington, DC 20361

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
N63733N	SP-01	0785-7P6	4780

11. TITLE *(Include Security Classification)*
The Effects of Asynchronous Visual Delays on Simulator Flight Performance and the Development of Simulator Sickness Symptomatology

12. PERSONAL AUTHOR(S)
K. C. Uliano, E. Y. Lambert, R. S. Kennedy, and D. J. Sheppard

13a. TYPE OF REPORT
Interim Final

13b. TIME COVERED
FROM 4/86 TO 12/86

14. DATE OF REPORT *(Year, Month, Day)*
26 December 1986

15. PAGE COUNT
74

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	08	
14	02	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Simulator sickness, visual delays, helicopter, flight simulation

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This research effort involved an experiment investigating the effect of asynchronous visual delays on simulator flight performance and the development of simulator sickness symptomatology. The SH-60B Vertical Takeoff and Land (VTOL) Simulator, part of the Navy's Visual Technology Research Simulator (VTRS) program was used to investigate this issue. ~~Three operational visual throughput~~ delays with varying amounts of asynchrony were tested: 215 ± 70 msec, 177 ± 23 msec, and 126 ± 17 msec. Twenty-five experienced pilots flew three 40- or 60-minute sessions with two simulator tasks (air taxi and slalom) under each of the lag conditions. Pilots flew one session per day for three days with the lag condition changing each day. Objective and self-report indices were collected and, while results showed no difference between lag conditions, paper-and-pencil illness ratings reflected a high initial incidence of illness (46% on Day 1) followed by rapid adaptation upon subsequent exposure. Simulator performance, however, was

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION

22a. NAME OF RESPONSIBLE INDIVIDUAL
Dr. Dennis C. Wightman

22b. TELEPHONE *(Include Area Code)*
(305) 646-5130

22c. OFFICE SYMBOL
N711

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. Abstract (cont'd)

differentially affected by lag with the longest lag producing the worst performance. Finally, relationships between sickness indices, flight performance data, and other variables are presented and discussed along with observations, recommendations, and areas for future research. Keywords: pilot studies; flight simulation.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

SECTION I

INTRODUCTION

BACKGROUND

The cost and training effectiveness of flight simulators has been attested to numerous times (e.g., Orlansky & String, 1977a, b). It is no secret that simulators will continue to train operational tasks. Furthermore, the present trend toward more extensive use of flight simulators will foster, due to the impact of increased costs, availability and safety issues of operating training aircraft (Caro, 1977). With emerging technological advances, the kinesthetic, vestibular, visual, and aircraft dynamic cues provided to the pilot in the flight simulator can closely approximate the real-world scenario. Consequently, advanced flight simulators such as those for training air combat maneuvering, formation flight, and air-to-ground weapons delivery are becoming more prevalent.

The addition of technological advances such as multiple computer image generation systems and high temporal and spatial resolution displays, aimed at increasing visual scene capabilities, are ultimately designed to increase the fidelity and realism of the flight simulator. Unfortunately, along with these technological advances, reports of simulator-induced distress have also increased and a collection of anecdotal and documented evidence of simulator sickness has begun to accumulate. (For an overview see Crosby & Kennedy, 1982; Ryan, Scott, & Browning, 1978; Money, 1980; and Kellogg, Castore, & Coward, 1980).

It is reasonable to assume that the extent to which the real system (e.g., an aircraft) produces motion sickness, a simulator which replicates the real environment is liable to induce the same responses. However, if sickness occurs in a simulator, but not in the real system, is some unique factor(s) or limitations of the simulator responsible for the induced sickness? For the purpose of this report, the term "simulator sickness" is reserved for those situations which are nauseogenic in the simulator but not in the corresponding aircraft.

The possible negative implications of simulator sickness can be grouped into three broad categories:

- Compromised Training. First, symptomatology may interfere with and retard learning in the simulator through distraction. Second, since humans are flexible, trainees may adapt to nauseogenic stimulation. If new learned processes are not similar to responses required in flight then the new responses could lead to negative or zero transfer to in-flight conditions. We believe this is a most critical problem because of the implications for safety of flight.

- <u>Decreased Simulator Use</u>. Due to the unpleasant side effects, simulators may not be used or persons may lack confidence in the training they receive in such devices.

- <u>Simulator Aftereffects</u>. Flying a simulator may result in aftereffects or posteffects. These are not unlike the posteffects of other motion devices, but their relevance to safety (e.g., egress from the simulator or driving home) is not known.

The evidence that exists suggests that simulator sickness symptomatology resembles motion sickness and other forms of distress which occur after exposure to altered and rearranged sensory information (Frank, Kennedy, McCauley, & Kellogg, 1983). Motion sickness is a general term for a constellation of symptoms and signs, generally adverse, brought on by exposure to abrupt, periodic, or unnatural accelerations. Overt manifestations are pallor, sweating, salivation, and vomiting (Kennedy & Graybiel, 1963a, b; Wiker, Kennedy, McCauley & Pepper, 1979a, b). Drowsiness, dizziness, and nausea are the chief self-report symptoms. Less frequently reported, but often present, are postural changes, or ataxia, sometimes referred to as "leans" or "staggers" (Fregly, 1974; Fregly & Kennedy, 1965). Other signs of motion sickness include changes in cardiovascular, respiratory, gastrointestinal, biochemical and temperature regulation functions (Colehour & Graybiel, 1966; McClure & Fregly, 1972; Money, 1970). Other symptoms include general discomfort, apathy, dejection, headache, stomach awareness, disorientation, lack of appetite, desire for fresh air, weakness, fatigue, confusion and, occasionally, incapacitation. Once symptoms become severe, time appears to be the only effective treatment.

The cue conflict theory (Reason, 1978) or sensory rearrangement theory (Steele, 1968) has generally been accepted as a working model for simulator sickness. Taken together, they purport that motion information from the visual, vestibular, and/or proprioceptive systems may be in conflict with expected values. These expected values are said to reflect past experience(s). McGuiness, Bouwman, and Forbes (1981) presented empirical evidence in support of this theory. They conducted an investigation of simulator sickness in the Navy's 2E6 Air Combat Maneuvering Simulator (ACMS). Twenty-seven percent of the aircrews using the ACMS reported varying degrees of symptoms. In addition, the more experienced aircrews (over 1500 flight hours) had a higher incidence of symptoms than less experienced flight crews.

CURRENT EMPHASIS

A program was begun within the Navy to study the simulator sickness problem. Field studies to document the problem of simulator sickness were conducted under the direction of the Human Factors Division of the Naval Training Systems Center (Kennedy, Dutton, Lilienthal, Ricard, & Frank, 1984; Kennedy, Dutton, Ricard, & Frank, 1984; Kennedy, Merkle, & Lilienthal, 1985; Lilienthal & Merkle, 1986; McCauley, 1984). Their findings are summarized below:

- Simulator sickness symptoms generally occur during simulator training but in a few cases are delayed for up to 10 hours.
- There is postsimulator readaptation and visual flashbacks have been reported.
- The overall incidence of simulator sickness can run as high as 50% or more if only initial hops are counted.
- Simulator sickness symptoms resemble those of more traditional forms of motion sickness except there appears to be a greater preponderance of headache and related visual epiphenomena.
- General discomfort, vertigo, and dizziness are the next most frequent complaints.
- Analysis of incidence rates ranged from 16%-55% for the simulators surveyed.

SECTION II

PURPOSE OF THE STUDY

Although the Navy has begun efforts to implement the fixes which emerged from a biomedical engineering panel review of simulator sickness (Berbaum, Kennedy, & Dunlap, 1986), interim "fixes" have dealt largely with changes which could be effected to the individual in order to promote more rapid adaptation to the devices. Insufficient attention has been paid to engineering criteria for new simulators and of retrofits to existing simulators. It has been found that pilots flying in simulators with particular asynchronous visual lags tend to develop symptoms traditionally associated with motion sickness. Thus, as a first step in such an enterprise, the experiment reported here was undertaken to isolate and study the relationship of visual asynchrony to simulator performance and to simulator sickness incidence.

Previous simulator design research has shown that lags in visual displays can be disruptive of performance (Ricard, Norman, & Collyer, 1976), and the disruption appears proportional to the duration of the lag. Very little research, however, has shown lags to be implicated in occasioning distress. An exception is the Miller and Goodson (1958) report where sickness prompted engineers to reevaluate the simulation. The sickness present was far greater than was to be expected in similar exposure in flight. Excessive lags were implicated as a possible factor contributing to the incidence of simulator sickness.

Most modern flight trainers employ computer image generation (CIG) visual displays. The nominal range of visual transport delays runs from 50 msec for simple systems to greater than 200 msec for more complex and/or poorly integrated CIGs. Conventional wisdom is that phase shifts of less than 30 degrees to 45 degrees at 1 Hz (83 - 125 msec) probably will not affect the control of a flight simulation (Ricard & Puig, 1977). The standard which governs motion platform systems (Department of Defense, 1974) proposes that visual asynchrony be avoided. Not taken into account is whether certain delays are more or less conducive to simulator sickness. It is not necessary that performance deficit and physical discomfort follow the same functional relationship relative to the magnitude of delay.

The work of Smith (1963) has shown that there are performance difficulties when information is visually delayed and motion related discomfort is implied. The magnitude of the delay which degrades motor performance may not be the same value (in msec) as the interval which one might find most distressing. Both of these forms of delay are present in flight simulators, but generally only the delay which intrudes on performance is studied. The latter is of importance for understanding simulator sickness. In general, the motor deficit is proportional to the magnitude of the visual delay, but delayed auditory feedback is most disturbing at about 100 msec (Rapin, Costa, Fromovitz, & Mandel, 1963). Observed effects of feedback delays indicate that little or no learning occurs in most response systems with feedback delays longer than 300 msec or, if limited learning occurs, it is likely to be unstable (Held, Efstathiou, & Greene, 1966). Indeed, visual delays as short

as 70 msec produced poorer performance (Baron, 1982). Similarly, Westra and Lintern (1985) investigated the effects of visual lags on simulated shipboard landing for a helicopter. They used delays of 117 and 217 msec and, although the effects were small in this particular study, poorer performance on vertical velocity and ship roll position at touchdown were associated with the longer lag. In addition, pilots rated the longer lag poorer in terms of fidelity and adequacy for training. These and other findings indicate that every motion system of the body is specialized in terms of the temporal feedback compliances that regulate it. If one accepts the conflict theory as an explanatory principle of motion sickness incidence, then it is possible that temporal rearrangements may produce an intersensory conflict analogous to those spatial distortions experienced with mirrors and prisms (cf., Welch, 1978, and Dolezal, 1982, for a review).

In a recent review of the 2F64C helicopter simulator at Jacksonville, Florida, Naval Air Station, a high incidence of simulator sickness was reported. At the same time, it was discovered that there were also large asynchronies between inertial and visual motions (Browder & Butrimas, 1983, 1984; Evans, Scott, & Pfeiffer, 1984). The size of these discrepancies was more than 200 msec (greater than what is typically proposed for such relationships between the moving platform and the computer generated imagery) and were not constant.

Another simulator (2F120 in Tustin, California) where similar visual asynchrony values have been reported (Browder & Butrimas, 1983, 1984; Evans, Scott, & Pfeiffer, 1984) was surveyed by Lilienthal and Merkle (1986). In that study, although delays were noted, they appeared to be smaller and less variable. Although it could not be determined whether the incidence of sickness was related to these delays, there were significant amounts of simulator sickness reported in both devices.

The present study was designed to investigate the relationship between asynchronous lags in the visual display and indicants of simulator sickness. Accordingly, the asynchronous lags present in the 2F64C and 2F120 simulators were modelled in the VTRS. These lags along with the VTRS minimum lag as a reference then constituted the primary conditions for an experiment in which pilots performed helicopter flight tasks. Various measures were taken during and after simulator flights to assess the effects of the lag conditions on performance and illness. Equally as important as the effects of lag were two additional areas of interest. First, and actually preceding the execution of this study, was establishing nauseogenic flight tasks which could serve for this experiment as well as for future investigations; and second, conducting measurement research on various indices of simulator sickness.

DEVELOPMENT OF THE EXPERIMENTAL MEASURES

MOTION SICKNESS SYMPTOMATOLOGY. The experimental assessment of motion sickness was probably first conducted by the Wendt group (Alexander, Cotzin, Hill, Ricciutti, & Wendt, 1945, a, b, c, d) during and immediately following World War II. Prior to that time nearly all studies employed vomiting as the criterion. Wendt's group developed a 3-point scale to assess motion sickness that was modified into a 5-point scale by Kennedy and Graybiel (1965). The

scale was later modified to a 9-point scale (Graybiel, Wood, Miller, & Cramer, 1968), and later a 7-point scale (Wiker, Kennedy, McCauley, & Pepper, 1979a). [illegible]ke, Homick, Ryan, and Mosely (1984) recently provided a 16 point scale. Regardless of the scaling properties, however, the criterion of motion sickness appears to be better measured with a questionnaire than by using emesis alone, perhaps because of the greater number of gradations available with symptomatology scoring (Wiker et al., 1979a). The theory behind motion sickness severity scaling is that vomiting, while the cardinal sign of motion sickness, is ordinarily preceded by a combination of symptoms (Lentz & Guedry, 1978). Thus, a modification of the questionnaires developed by Kennedy, Dutton, Ricard, and Frank (1984) were used in this experiment to assess motion sickness.

MOTION HISTORY QUESTIONNAIRE. A series of studies have been conducted with an anamnestic form which inquires into a subject's history and exposures to different motion environments. Scores on this test have been related to success in flight training (Hutchins & Kennedy, 1965) and it was shown that subjects respond truthfully (Hardacre & Kennedy, 1963). However, only modest relationships have been found in prediction of simulator sickness incidence (Kennedy, Frank, McCauley, Bittner, Root, & Binks, 1984).

POSTURAL EQUILIBRIUM. Motion-induced vestibular ataxia is not widely known, but has been reported following protracted exposures to a centrifuge and to ships at sea (Fregly, 1974). Both postural equilibrium and tracking are closed-loop psychomotor control systems under voluntary guidance by the cerebral cortex and under automatic (i.e., motor) control in the cerebellum (Hill, 1971; Stockwell, Koozekanani, & Barin, 1981). Thus, it is reasonable to expect that if posture is disrupted by exposure to motion, human manual control may be similarly affected. The motion-induced ataxia may be analogous to eye-hand coordination changes following rearranged visual feedback which occurs when wearing reversing or displacing prisms (Welch, 1978). A short procedure for measuring postural equilibrium is available (Thomley, Kennedy, & Bittner, 1986) which correlates well with the full scale battery (Fregly, Smith, & Graybiel, 1973; Graybiel & Fregly, 1965).

GRAMMATICAL REASONING. An Automated Portable Test System (APTS) is currently under development that will contain a series of cognitive and motor tasks which have been shown to have excellent metric properties (Bittner, Smith, Kennedy, Staley, & Harbeson, 1985; Kennedy, Wilkes, Lane, & Homick, 1985). The battery is implemented on a personal computer (NEC PC8201A) and has been tested in different populations and environments. Thus far, one of the tests most sensitive to treatment effects is the Grammatical Reasoning test of Baddeley (1968). Previous studies show the test to have very high test-retest reliability and stability (Carter & Kennedy & Bittner, 1981). In addition, Schifflett, Bowes, and Haswell (1985) and Coussens (1985) have reported reductions in performance on the Grammatical Reasoning test with hypoxia. The Grammatical Reasoning test takes descriptive statements of letter arrangements (e.g., "A comes before B") and invites true/false comparison responses with the real relationships (e.g., "BA"). Sixteen statements are employed (half postive and half negative; half are true and half are false) and these are repeated in random fashion for a period of 90 seconds. Subsequent administrations presented different combinations of the questions to the subject.

scale was later modified to a 9-point scale (Graybiel, Wood, Miller, & Cramer, 1968), and later a 7-point scale (Wiker, Kennedy, McCauley, & Pepper, 1979a). Reschke, Homick, Ryan, and Mosely (1984) recently provided a 16-point scale. Regardless of the scaling properties, however, the criterion of motion sickness appears to be better measured with a questionnaire than by using emesis alone, perhaps because of the greater number of gradations available with symptomatology scoring (Wiker et al., 1979a). The theory behind motion sickness severity scaling is that vomiting, while the cardinal sign of motion sickness, is ordinarily preceded by a combination of symptoms (Lentz & Guedry, 1978). Thus, a modification of the questionnaires developed by Kennedy, Dutton, Ricard, and Frank (1984) were used in this experiment to assess motion sickness.

MOTION HISTORY QUESTIONNAIRE. A series of studies have been conducted with an anamnestic form which inquires into a subject's history and exposures to different motion environments. Scores on this test have been related to success in flight training (Hutchins & Kennedy, 1965) and it was shown that subjects respond truthfully (Hardacre & Kennedy, 1963). However, only modest relationships have been found in prediction of simulator sickness incidence (Kennedy, Frank, McCauley, Bittner, Root, & Binks, 1984).

POSTURAL EQUILIBRIUM. Motion-induced vestibular ataxia is not widely known, but has been reported following protracted exposures to a centrifuge and to ships at sea (Fregly, 1974). Both postural equilibrium and tracking are closed-loop psychomotor control systems under voluntary guidance by the cerebral cortex and under automatic (i.e., motor) control in the cerebellum (Hill, 1971; Stockwell, Koozekanani, & Barin, 1981). Thus, it is reasonable to expect that if posture is disrupted by exposure to motion, human manual control may be similarly affected. The motion-induced ataxia may be analogous to eye-hand coordination changes following rearranged visual feedback which occurs when wearing reversing or displacing prisms (Welch, 1978). A short procedure for measuring postural equilibrium is available (Thomley, Kennedy, & Bittner, 1986) which correlates well with the full scale battery (Fregly, Smith, & Graybiel, 1973; Graybiel & Fregly, 1965).

GRAMMATICAL REASONING. An Automated Portable Test System (APTS) is currently under development that will contain a series of cognitive and motor tasks which have been shown to have excellent metric properties (Bittner, Smith, Kennedy, Staley, & Harbeson, 1985; Kennedy, Wilkes, Lane, & Homick, 1985). The battery is implemented on a personal computer (NEC PC8201A) and has been tested in different populations and environments. Thus far, one of the tests most sensitive to treatment effects is the Grammatical Reasoning test of Baddeley (1968). Previous studies show the test to have very high test-retest reliability and stability (Carter & Kennedy & Bittner, 1981). In addition, Schifflett, Bowes, and Haswell (1985) and Coussens (1985) have reported reductions in performance on the Grammatical Reasoning test with hypoxia. The Grammatical Reasoning test takes descriptive statements of letter arrangements (e.g., "A comes before B") and invites true/false comparison responses with the real relationships (e.g., "BA"). Sixteen statements are employed (half postive and half negative; half are true and half are false) and these are repeated in random fashion for a period of 90 seconds. Subsequent administrations presented different combinations of the questions to the subject.

SECTION III

METHOD

SUBJECTS

Twenty-five experienced Navy and Marine helicopter pilots participated in the experiment. All pilots were male volunteers from operational flight squadrons. The pilots averaged 1,071 hours of total flight time and varied in overall experience with a range of 360 to 2850 flight hours. Pilots averaged 66.68 hours of simulator experience with a range of 4 to 250 hours. Table 1 provides a biographical outline of each pilot; Table 2 further summarizes the subjects' overall aviation experience. Subject 21 arrived ill at the VTRS and, although his data were collected, they were removed from subsequent analyses.

TABLE 1. BIOGRAPHICAL DATA OF PILOTS

S	Total Flight Hrs	Sim Hrs	Last Sim Hop	Sim Type	A/C Type
1	630	50	20 mos.	2F64C	SH-3H
2	850	50	9 mos.	2F135	SH60B
3	700	50	N/A	N/A	N/A
4	800	100	12 mos.	2F106	SH-2F
5	900	100	8 mos.	2F106	SH-2F
6	600	30	10 mos.	2F64C	SH-3H
7	630	75	10 mos.	2F64C	SH-3H
8	1210	50	7 mos.	2F135	SH60B
9	475	50	8.5 mos.	2F135	SH60B
10	450	200	8 mos.	2F106	SH-2F
11	800	100	10 mos.	2F106	SH-2F
12	1130	45	2 wks.	2F117	CH46E
13	1300	110	5 days	2F121	CH53D
14	600	250	5 days	2F120	CH53E
15	2600	60	2 mos.	2F120	CH53E
16	2500	4	1 mo.	2F121	CH53D
17	1500	30	14 days	2F117	CH46E
18	2800	4	N/A	2B24	UH-60
19	1251	20.3	4 yrs.	2B24	UH-60
20	1075	33.6	4 yrs.	2B24	H-1
21*	920	20	2 yrs. 7 mo.	2B24	H-1
22	470	50	2 wks.	2F120	CH53E
23	2850	10	1 mo.	2F117A	CH46E
24	1400	50	32 days	2F120	CH53E
25	360	125	3 wks.	2F120	CH53E

* Removed from analyses
Mean Flight Hours = 1071.24
Mean Simulator Hours = 66.68
N/A = Not available

TABLE 2. AVIATION EXPERIENCE

	Flight Time	Simulator Time
Mean	1071	66.68
SD	27.39	7.64
Range	360-2850	4-250
N	24	24

APPARATUS

The Vertical Take-off and Landing (VTOL) Simulator at the Visual Technology Research Simulator (VTRS) consists of a cockpit which is representative of the Navy's SH-60B Seahawk helicopter, a motion seat, and a wide-angle visual system (Herndon, 1982). The cockpit is mounted on top of a fixed platform and enclosed in a spherical (34 feet diameter) dome. The cockpit is provided with instrumentation and controls for the right seat. All basic aircraft systems are simulated with limited navigation and emergency procedures available. The motion seat was not used for this experiment.

The visual scene was represented by computer-generated images that are projected onto a 34-foot diameter Spitz dome. A General Electric Compu-Scene I (upgraded to an extra edge capacity of a Compu-Scene III) and a PDP 11/55 minicomputer were used to provide a 4000-edge capacity. Two full-color TV light valve projectors (1025 lines) were used to display the imagery in adjacent fields to give a 160 degree horizontal (40 degrees left and 120 degrees right) by 70 degrees vertical (20 degrees up to 50 degrees down) field of view. Brightness and contrast levels of the projectors were left constant throughout the experiment.

Aerodynamic and visual subsystem computations were computed at a 30 Hz iteration rate by a SEL 32/77 minicomputer system which is comprised of three high speed multiple processors. Cyclic, collective, and directional pedal control loading was provided by a McFadden variable force control loading system. Aircraft and environmental sounds were also simulated.

EXPERIMENTAL FACTORS

Each subject flew two experimental tasks under three separate visual transport delays (i.e., the lag between a pilot's input sample and the first frame of a video output corresponding to that sample) which were administered across three days. The three lag conditions simulated were from three different Navy simulators. (1) SH3 (Device 2F64C) condition with a lag of 215 ± 70 msec; (2) CH53E (Device 2F120) condition with a lag of 177 ± 23 msec; and (3) VTRS standard condition with a lag of 126 ± 17 msec. Table 3 presents the helicopter and corresponding simulator designations.

TABLE 3. HELICOPTER AND SIMULATOR DESIGNATIONS

Simulated Helicopter		Simulator Designation	Lag (msec)	
			Mean	Range
SH-3	Sea King	2F64C	215	±70
CH53E	Super Stallion	2F120	177	±23
SH60B	Seahawk	VTRS	126	±17

TASKS

The experimental tasks were an air taxi maneuver around a helipad and a slalom maneuver down a runway. Extensive pre-experimental work found that the two tasks met the primary criteria for inclusion as tasks in the experiment. That is, both tasks were designed such that pilots must focus most of their attention on the visual scene outside the cockpit. Also, the tasks were successful in provoking symptoms related to motion sickness. The air taxi and slalom tasks are described below.

AIR TAXI. Pilots began the task on a corner of the helipad. They were required to achieve 25 feet of altitude and proceed on a cardinal heading while keeping the edge of the pad just inside their chin bubble. Pilots were also instructed to keep the position of the aircraft's nose coincident with the pad at each corner while performing a pivot (or pedal) turn. At the initial starting point, the pilots were told to hover for 30 seconds then repeat the maneuver. The experimenter instructed the pilots to stop after either 20 or 30 minutes depending on the experimental condition. Figure 1 shows the helipad surrounded by the other objects contained in the data base.

SLALOM. Pilots began the task 1/4 mile from a runway on the ground. They were required to achieve 100 feet of altitude, 70 knots of airspeed and fly a slalom course down the runway reversing bank angle at each one of the 12 hash marks while crossing through the centerline. For the purpose of this task, lateral movement was confined to the 200 feet width of the runway, assuming that the pilots performed the task as specified. This task was generally accomplished between 45-60 seconds and would thereby translate into an oscillation frequency between 0.10-0.23 Hz -- a very nauseogenic range for linear motions (McCauley & Kennedy, 1976). At the end of the runway, the pilots were instructed to maintain altitude and perform a standard right-hand turn, heading downwind and proceed back to the initial starting point. The experimenter instructed the pilot to stop after either 20 or 30 minutes depending on the experimental condition. Figure 2 shows an overhead view of the runway, while Figure 3 demonstrates a 30-degree bank to the right.

EXPERIMENTAL DESIGN

A repeated-measures design with each pilot flying both the slalom and air taxi tasks under all three visual lag conditions. Pilots flew one session consisting of both tasks and one lag condition on each of three consecutive days. The lag condition was varied across the three days for each pilot; tasks and lags were fully counterbalanced across both pilots and days. The experimental design is summarized in Table 4.

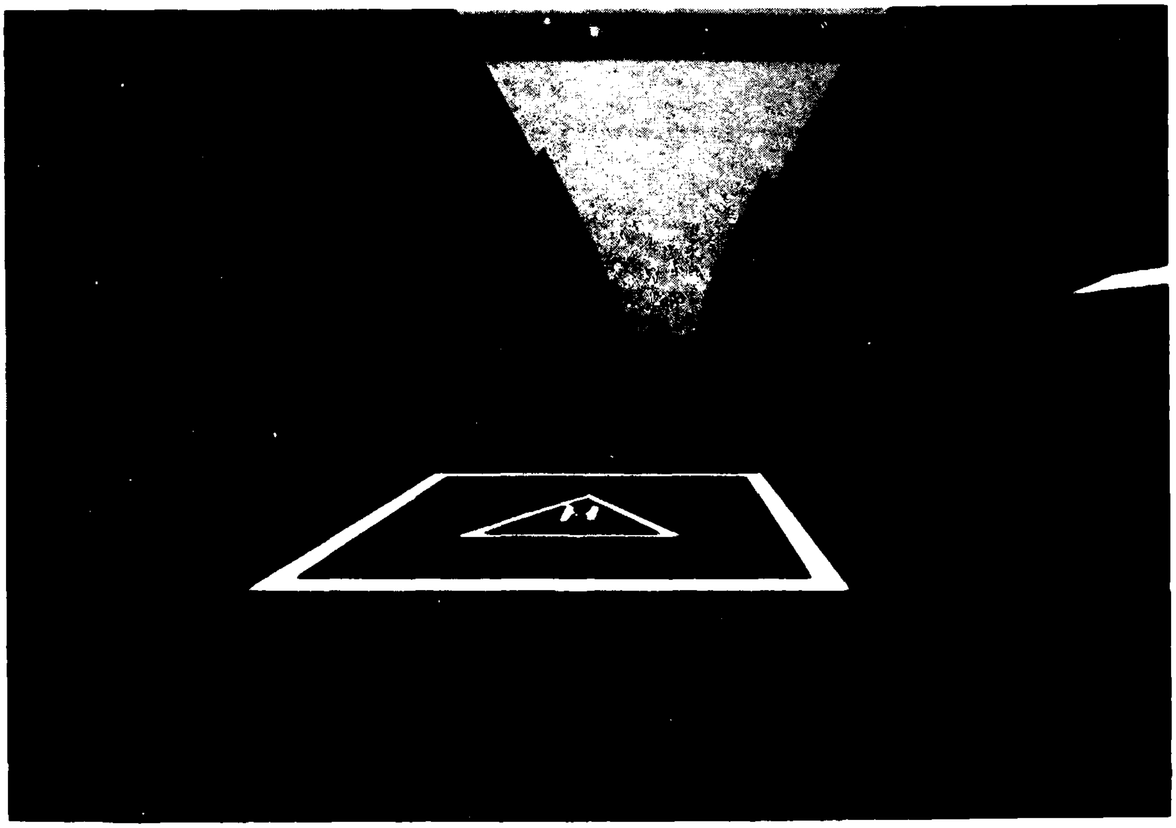


Figure 1. View of the Helipad and Other Data Base Components Used for the Air Taxi Task

Figure 2. Overhead View of the Runway and Data Base Components Used for the Slalom Task

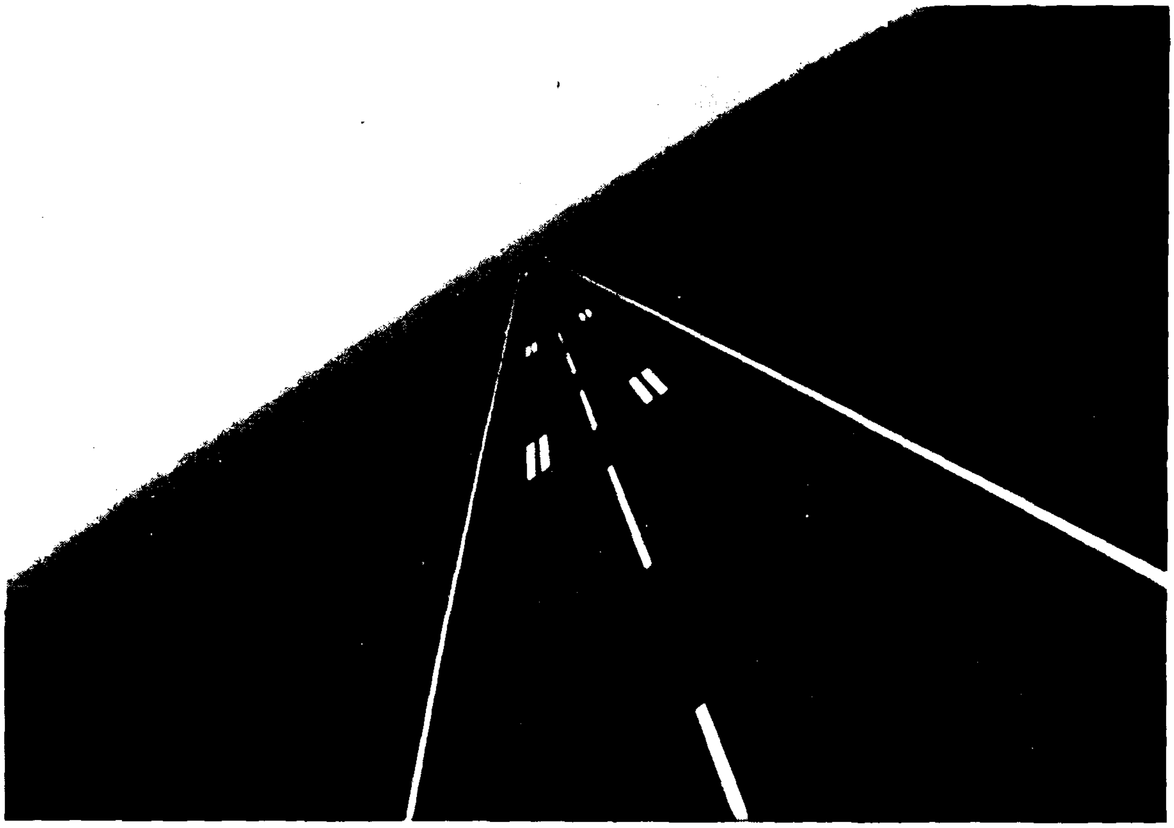


Figure 3. View of a Simulated 30-Degree Bank Angle
While Flying a Slaom Course

TABLE 4. EXPERIMENTAL DESIGN

Pilot*	Day 1	Day 2	Day 3
1, 12	A I B	B II A	A III B
2, 11	B II A	A III B	B I A
3, 10	A III B	B I A	A II B
4, 21	A I B	B III A	A II B
5, 20	B II A	A I B	B III A
6, 19	A III B	B II A	A I B
7, 18	B III A	A II B	B I A
8, 17	A II B	B I A	A III B
9, 16	B I A	A III B	B II A
23, 14	B I A	A II B	B III A
22, 13	A II B	B III A	A I B
24, 15	B III A	A I B	B II A

Tasks: A = Air Taxi
B = Slalom

Lags: I = 215 ± 70 msec
II = 177 ± 23 msec
III = 126 ± 17 msec

*Pilots 1-12 flew each task for 20 minutes
Pilots 13-24 flew each task for 30 minutes

The experiment was designed so that pilots performed each task for 20 minutes for a total of 40 minutes of flight time each day under a specific display lag. However, if the pilots experienced physical discomfort and requested to exit the simulator prior to completion of 40 minutes flight time, their hop was considered terminated for the day. In either case, sickness and simulator performance data were collected.

Preliminary analysis of the data midway through the experiment indicated that a smaller than expected percentage of the pilots were reporting symptoms of simulator sickness. Thus, in an attempt to tap into more symptoms and incidence of simulator sickness and increase the power of the experiment, some adjustments were made. First, total flight time in the simulator was increased from 40 to 60 minutes (i.e., from 20 to 30 minutes per task). Second, a small amount of turbulence and 5 knots of wind were inserted in the air taxi maneuver. These changes were made such that the orthogonal integrity of the design was preserved.

SIMULATOR PERFORMANCE MEASUREMENT

Aircraft altitude, latitudinal, and longitudinal positions were sampled at 30 Hz for the slalom and air taxi tasks and constituted the raw data. More specifically, the longitudinal measure along with the latitudinal measure, calculated for the air taxi task were used to compute lateral deviation from specified task performance in the X-Y coordinate system. Therefore, two lateral measures of deviation for the air taxi task are reported; two "X" legs of the helipad square (South to North and North to South); and two "Y" legs

(West to East and East to West) (refer to Figure 1). Similarly, the longitudinal and latitudinal measures for the slalom task were used in computing the lateral deviation from the confines of the runway. In addition, airspeed was relevant and therefore sampled only for the slalom task. Finally, for the slalom task every time the aircraft crossed a set of hash marks while crossing through the centerline, a "hit" was recorded and a percentage of hits to misses was calculated.

Summary measures, calculated using the raw data described above, were percentage time-out-of-tolerance (TOT) scores and were calculated for each trial. For the air taxi maneuver the tolerance band defining acceptable performance was set at ±5 feet of altitude and ±2 feet for each lateral measure. For the slalom task, the tolerance bands were set at ±25 feet of altitude and ±5 knots of airspeed.

SIMULATOR SICKNESS MEASUREMENT

GLOBAL ILLNESS RATING. A paper-and-pencil rating scale was used separately by the experimenters and subjects to arrive at an index of the overall physical appearance of the subject. The index consisted of a 7-point anchored scale that was administered pre- and posthop on all three days. A post-minus prehop composite score was calculated. A copy of the rating scale is presented in Appendix A.

The illness rating procedure involved the subject rating himself prior to entering the simulator. Since the construction of the scale utilized the word "appearance," a mirror was provided to aid in the rating. In addition to the self-report measure, two experimenters were present and used an identical scale to independently rate the subject immediately before and after the hop.

The composite illness score for the experimenters was calculated by using the mean difference between the post-minus prehop ratings for the two experimenters.

TASK DIFFICULTY SCALE. A paper-and-pencil rating scale queried the pilots regarding the difficulty of the task. Operationally, difficulty was defined in terms of the degree of attention and effort that each task required of the pilot. The index consisted of a 7-point anchored scale that was administered pre- and posthop on all three days. A copy of the rating scale is presented in Appendix B.

SUBJECT PHOTOGRAPHS. A Pentex K1000 SLR 35mm tripod-mounted camera with a 80-205 Soligar macrolens was used to take pre- and posthop photographs of the subject on all three days. The camera was set at a shutter speed of 1/60th second along with an aperture setting of 8. A standard strobe-type flash was used to counteract the effect of the fluorescent lighting present in the building. Kodak VR (ASA 100) film was used and professionally developed, not allowing for color or contrast corrections. The final prints measured 9cm x 12.5cm.

The photographic procedure involved the subject standing approximately 3m from the camera facing forward. A black backdrop provided the background. If the subject wore eyeglasses, he was asked to remove them and look into the

camera lens without any facial expressions. The subject was focused in such a manner as to produce a front bust view.

Six naive raters, using the global illness rating scale, rated the photographs (a description is provided in Appendix C). These scores were subsequently compared to on-site ratings provided by the pilots as well as the two experimenters.

POSTURAL EQUILIBRIUM TEST. Two Postural Equilibrium tests were used to assess ataxia as a sign/symptom of simulator sickness. The two tests, Stand-on-Preferred-Leg (SOPL) and Stand-on-Nonpreferred-Leg (SONPL), are modified versions of the Fregly-Graybiel Battery (Fregly, Smith, & Graybiel, 1973) and were shown to be effective, stable and reliable measures of ataxia (Thomley, Kennedy, & Bittner, 1986).

Subjects were asked to stand on their "preferred" leg in an upright position with arms folded across their chest and eyes closed for a maximum of 30 seconds. The experimenter used a stopwatch to time how long the subject stood without sidestepping, losing his balance, or otherwise deviating from the position. The trial ended either at the 30 second maximum limit or when a deviation from position occurred. Each subject performed the SOPL test for three consecutive trials. The SONPL test was administered in the same manner as SOPL, within the same criteria, except the subject stood on his chosen "nonpreferred" leg. Both tests were administered to each subject, pre- and posthop, for all three days.

GRAMMATICAL REASONING TEST. A performance test was self-administered pre- and posthop on each of the three days. The paper-and-pencil based Grammatical Reasoning test was chosen from a battery developed under the Performance Evaluation Test for Environmental Research program (PETER) (Bittner, Carter, Kennedy, Harbeson, & Krause, 1986). The test was later mechanized on a portable microprocessor called the Automated Portable Test System (APTS) (Bittner, Smith, Kennedy, Staley, & Harbeson, 1984). The APTS uses a NEC PC 8201A, an 8-bit device configured around an 80C85 microprocessor and 32K internal read-only memory containing BASIC. The entire package measures 9" x 12" x 2.5," is battery operated and easily transportable. The NEC PC was used to administer the test in this experiment.

The Grammatical Reasoning test purports to assess effects on analytic-cognitive functioning. A statement such as "A does not procede B" was presented, followed by a pair of letters, either "AB" or "BA." The subject depressed the "T" or "F" key depending on whether the statement describing the pair of letters was true or false. Administration time was under two minutes and the number of administrations varied between 17 and 46 trials. Mean response latency and percentage correct were the dependent measures of interest across the three days.

MOTION SICKNESS QUESTIONNAIRE (MSQ). A self-report questionnaire, the Motion Symptomatology Questionnaire, was administered to each pilot, pre- and posthop, on all three days. The prehop portion contained questions pertaining to the pilot's present physiological status and a motion sickness symptom checklist. The same symptom checklist was used to gather posthop information.

A response was required for each symptom using a rating of "none," "slight," "moderate," or "severe," or a "yes/no" for some questions. The 27 symptoms were subsequently grouped according to a Modified Diagnostic Categorization Score Sheet (Kennedy, Dutton, Lilienthal, Ricard, & Frank, 1984) as either pathogonomic (only vomiting), or symptoms classified as either *major*, *minor*, visual, *mental*, or other. A copy of the MSQ is presented in Appendix D.

The "Diagnostic Criteria for Levels of Motion Sickness Severity," a 7-point scaling method used in the Kennedy, Dutton, Ricard, and Frank (1984) study was used to score the pre- and postraw symptomatology data. A post-minus prehop composite score was calculated.

MOTION HISTORY QUESTIONNAIRE (MHQ). Pilots completed a paper-and-pencil form questionnaire, the MHQ, as part of the prehop measures on the first day only. The MHQ inquired into the pilot's recent flying and simulator experience, exposure and preference for motion devices, total flight hours, and information surrounding previous nauseating experiences. The questionnaire was used to evaluate the pilot's historical susceptibility to motion sickness. A copy of the MHQ is presented in Appendix E.

For review, Table 5 presents a summary of the measures used, when they were administered, and what they purport to assess.

PROCEDURE

Upon arrival at the VTRS laboratory, subjects were given an information packet which consisted of an experimental brief and informed consent material. This packet served to inform the subject of the nature of the experiment, procedures and measures used, and the equipment features of the VTOL simulator. Subjects were required to read and sign the consent form and a copy of the form was placed in their permanent health records.

Each pilot was assigned to one of the experimental conditions. Pilots completed their prehop measures and were cycled through their simulator session. The pilot was the only person in the cockpit during the hop and was in direct communication via a headset with an experimenter located at the Instructor/Operator Station. Instructional feedback was given during each trial if the pilot exceeded the tolerance bands for each maneuver for more than 10 seconds (e.g., "...your altitude is a little high"). Posthop measures were taken at either the end of their simulator session or when they exited the simulator due to illness. The pilots flew three consecutive days before returning to their squadrons. Discrete measurements were taken prehop, within the simulator, and posthop.

PREHOP. First, the MHQ and prehop MSQ were completed. Second, the subjects were given the Grammatical Reasoning test preceded by paper-and-pencil practice problems designed to insure understanding of the concepts and semantics underlying the task. This practice opportunity was administered only during the first day. The pilots were subsequently taken into an adjacent office and administered the Postural Equilibrium test. This was followed by the subject completing the global illness rating scale and finally the prehop photograph.

TABLE 5. SUMMARY TABLE OF PERFORMANCE MEASURES

Dependent Variable	Administered	Assesses
Global Illness Scale	Pre-/Posthop	Illness
Motion History Questionnaire (MHQ)	Prehop	Pilot susceptibility
Motion Sickness Questionnaire (MSQ)	Pre-/Posthop	Illness
Grammatical Reasoning	Pre-/Posthop	Analytic/Cognitive Functioning
Postural Equilibrium	Pre-/Posthop	Ataxia
Task Difficulty Scale	Posthop	Task Difficulty
Simulator Performance Measures (TOT)		
Slalom Task		
Lateral (width of runway)		
Airspeed		
Altitude	During hop	Flight Performance
Hit Rate (%)		
Air Taxi Task		
(X) Lateral (S-N, N-S)		
(Y) Lateral (W-E, E-W)	During hop	Flight Performance
Altitude		

IN-SIMULATOR. Pilots flew a 10-minute familiarization period (5 minutes per visual data base) in the simulator (day one only) prior to beginning their experimental sequence. During this time, they were free to practice the tasks if they chose. They were then set on the first task initial flight condition. The pilots performed the first maneuver (either slalom or air taxi). As described previously, pilots 1-12 flew 20-minute sessions, while pilots 13-24 flew 30-minute sessions. They were then asked to recall the global illness scale and give an intrahop verbal rating based on that scale.

The pilots were subsequently set up on the second task initial flight condition. They flew the second maneuver for either 20 or 30 minutes. When the session was completed, the pilot was ushered out of the simulator and the posthop procedure was initiated.

POSTHOP. The posthop measures involved, in order, the posthop photograph, global illness rating scale, Postural Equilibrium test, posthop MSQ, and finally the Grammatical Reasoning test. The procedure was administered in the same manner for the three day duration of the experiment.

SECTION IV

RESULTS

SICKNESS INDICES ANALYSES

Data were analyzed using analysis-of-variance (ANOVA) procedures. Lag had no effect on the paper-and-pencil sickness indices (i.e., pilot and experimenter provided illness composite ratings and the MSQ composite rating). The day main effect, however, was significant for the sickness indices with Tables 6-9 providing summary ANOVA tables. Specifically, testing day had an effect on the pilot's self-report composite illness rating ($F(2,40)=5.60$, $p < .025$), mean experimenter composite illness rating ($F(2,40) = 4.21$, $p < .05$), and MSQ composite rating ($F(2,40) = p < .10$). Moreover, testing day also affected the intrahop illness rating (i.e., the way in which the pilots rated their sickness at the midpoint of their hop, $F(2,40) = 10.08$, $p < .01$). Figure 4 presents the four sickness composite indices across the three testing days. As is evident, the effects seem to contain a strong linear component and, upon examination, the magnitude of these indices did indeed decrease significantly across each day. Finally, there were neither significant day x lag interactions nor statistically reliable differences in equilibrium or cognitive/analytical functioning, as measured by the Postural Equilibrium and Grammatical Reasoning tests respectively as functions of lag or day.

In an effort to test the effectiveness of the cue conflict theory as a working model for simulator sickness, separate analyses were performed using two, one-way ANOVAs. One analysis was conducted with flight experience as the grouping variable; the other with simulator experience as the grouping variable. These variables were categorized into subgroups and, in an effort to make the number in each subgroup as similar as possible, the following cutpoints were assigned: 0-899 flight hours were considered low ($n=12$); 900 and above were considered high ($n=12$). Likewise, 0-50 simulator hours were considered low ($n=15$); 51 and above was considered high ($n=9$). This arrangement produced a matrix containing four cells: low-low, low-high, high-low, and high-high. Analyses revealed that the amount of previous flight and simulator experience was related to the pilot's self-report composite illness rating ($F=(1,70) = 14.29$, $p < .001$), mean experimenter composite illness rating ($F(1,70) = 15.76$, $p < .001$), and MSQ composite rating ($F(1,70) = 6.60$, $p < .01$). It appears that both more flight experience and less simulator experience are associated with greater illness as a result of simulator exposure; however, as indicated by the previous analysis, this illness quickly dissipates with time.

As described earlier, pilots rated how much attention and effort each task required of them on each day using a 1 to 7 anchored scale. In general, pilots found both tasks to be moderately difficult with marginal means of 4.83 for slalom and 4.90 for air taxi. Analysis of task ratings are presented in Tables 10 and 11. The slalom task yielded no lag effect, yet a strong day effect ($F(2,40) = 9.31$, $p < .01$) revealing that the task was perceived as being less difficult with each day. The significant day x lag interaction represented graphically in Figure 5 was a result of the VTRS cofiguration

TABLE 6. ANALYSIS OF VARIANCE FOR THE MOTION SICKNESS QUESTIONNAIRE COMPOSITE RATINGS

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	27.08	2	13.54	2.51	.10
Lag (L)	1.08	2	0.54	0.10	
D x L	1.17	4	0.29	0.05	
Pilots	118.88	23	5.17	0.96	
Residual	214.54	40	5.39		
TOTAL	363.75	71			

TABLE 7. ANALYSIS OF VARIANCE FOR THE INTRAHOP ILLNESS RATING

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	32.86	2	16.43	10.08	.01
Lag (L)	2.11	2	1.06	0.65	
D x L	1.18	4	0.30	0.18	
Pilots	37.99	23	1.65	1.01	
Residual	65.17	40	1.63		
TOTAL	139.31	71			

TABLE 8. ANALYSIS OF VARIANCE FOR THE
SELF-REPORT COMPOSITE ILLNESS RATING

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	32.86	2	16.43	5.60	.025
Lag (L)	6.03	2	3.02	1.03	
D x L	2.15	4	0.54	0.18	
Pilots	83.78	23	3.64	1.24	
Residual	117.41	40	2.94		
TOTAL	242.23	71			

TABLE 9. ANALYSIS OF VARIANCE FOR THE MEAN
EXPERIMENTER COMPOSITE ILLNESS RATING

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	22.22	2	11.11	4.21	.05
Lag (L)	0.34	2	0.17	0.06	
D x L	5.59	4	1.40	0.53	
Pilots	76.99	23	2.64		
Residual	105.67	40			
TOTAL	210.75	71			

TABLE 10. ANALYSIS OF VARIANCE FOR SLALOM TASK DIFFICULTY RATINGS

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	15.64	2	7.82	9.31	.01
Lag (L)	0.23	2	0.12	0.14	
D x L	11.72	4	2.93	3.49	.05
Pilots	136.00	23	5.91	7.04	.01
Residual	33.51	40	0.84		
TOTAL	197.10	71			

TABLE 11. ANALYSIS OF VARIANCE FOR AIR TAXI TASK DIFFICULTY RATINGS

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	26.86	2	13.43	55.96	.01
Lag (L)	2.56	2	1.28	5.33	.01
D x L	22.40	4	5.60	23.33	.01
Pilots	114.99	23	5.00	20.83	.01
Residual	9.51	40	0.24		
TOTAL	176.32	71			

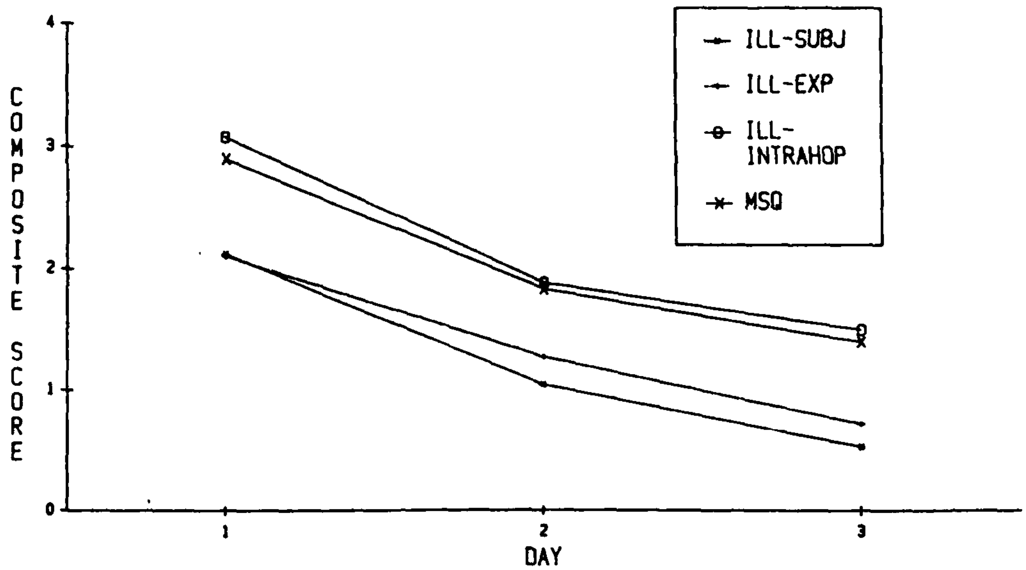


Figure 4. Sickness Composite Scores Across Days

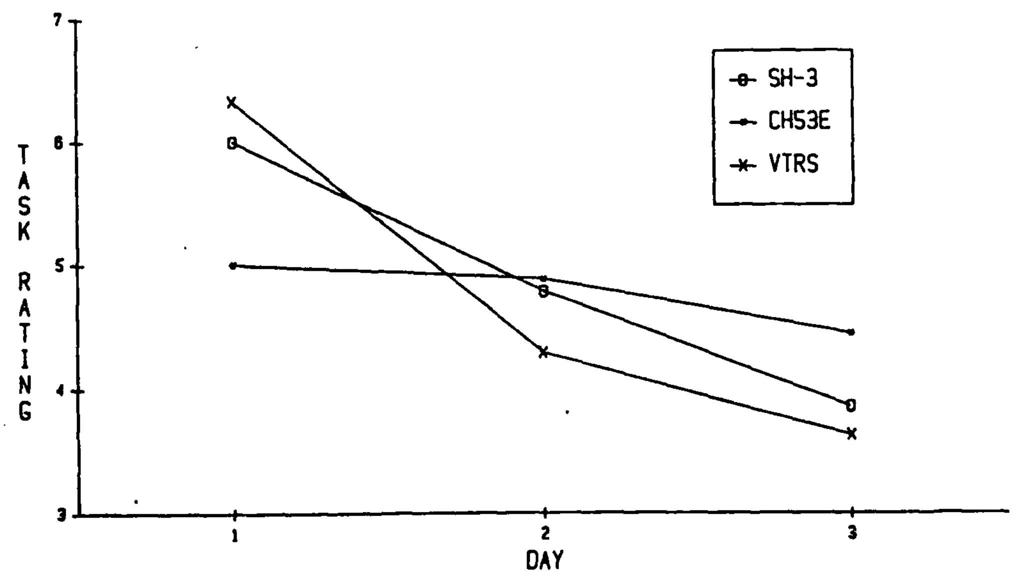


Figure 5. Mean Task Difficulty Rating for the Slalom Task Across Days

having the highest difficulty rating on Day 1, yet subsequent ratings on Days 2 and 3 were the lowest compared to the CH-53E and SH-3 conditions. The air taxi task difficulty demonstrated a strong day effect ($F(2,40) = 55.96$, $p < .001$), as well as a significant lag effect ($F(2,40) = 5.33$, $p < .01$). Post hoc comparisons showed the day effect, as predicted, to be linear in nature. Also, for the lag effect, the CH-53E condition (moderate lag, moderate variability) was perceived as being more difficult than either the SH-3 or VTRS conditions with no difference between these latter two. These main effect findings must be evaluated against the significant day x lag interactions which indicated that the air taxi task was perceived as being easier across Days 1 through 3 for the SH-3 and CH-53E conditions; however, for the VTRS condition, a decrease in difficulty from Day 1 to Day 2 is followed by an increase from Day 2 to Day 3 (see Figure 6).

Table 12 presents a portion of the correlation matrix among sickness indices and category groupings, simulator performance data, and task difficulty ratings. The intercorrelations between the illness and MSQ composite ratings were significant, ranging from 0.58 to 0.71, indicating extensive measurement overlap and providing an independent verification of the reliability of these measures. The fact that the intrahop illness ratings are associated with end-of-hop illness ratings indicate that sickness ratings at either 20 or 30 minutes into the hop are similar to those at the hop's conclusion (i.e., either 20 or 30 minutes later).

The four sickness indices were also each significantly related to the two equilibrium tests indicating that, as the degree of pilot and experimenter-reported illness increased, pilots' ability to maintain postural equilibrium decreased. Pilot and experimenter-provided composite illness ratings were related ($r = .71$, $p < .01$) demonstrating strong experimenter/pilot agreement. In addition, comparison of the two experimenter's composite ratings yielded an interscorer reliability index of 0.87. There was also a moderately strong relationship between task difficulty ratings and sickness indices. Evidently, the more difficult the pilots perceived the task, the greater the degree of illness.

Evaluation of the ratings of photographs by the naive raters showed no significant difference between the composite scores given by the pilots, onsite experimenters, and naive raters. Furthermore, the illness composite scores for the pilots and raters were related ($r = .45$, $p < .01$) as were the composite scores for the raters and experimenters ($r = .59$, $p < .01$). The mean interscorer reliability index for the naive raters was 0.67 with a range of 0.51 to 0.77. This estimate of scorer reliability was considerably less than that value obtained for the onsite experimenters which was 0.87.

The relationship between the amount of flight and simulator experience and the sickness indices do provide further support for the cue conflict theory of simulator sickness. The amount of flight hours was positively related to each of the sickness indices, while the amount of simulator experience was negatively related to the same sickness indices.

A final observation from Table 12 compares the task difficulty rating for each task with simulator flight performance data corresponding to that task. It would be expected that the more difficult the task is perceived the poorer the performance would be. This was true for the air taxi task but generally not for the slalom task.

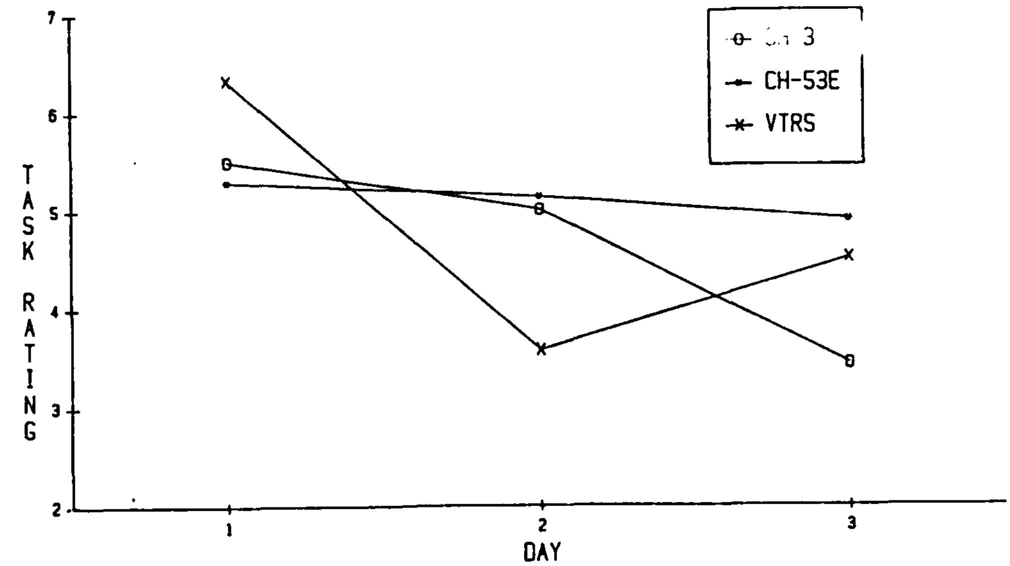


Figure 6. Mean Task Difficulty Rating for the Air Taxi Task

SIMULATOR PERFORMANCE DATA ANALYSIS

AIR TAXI. Tables 13-15 present summary ANOVA tables for lateral (X, Y) and altitude TOT summary measures respectively. Lag had no effect on any of these three measures. The performance on the X and Y lateral TOT measures, however, did show a significant day effect ($F(2,40) = 9.75$, $p < .001$), $F(2,40) = 8.54$, $p < .001$). Further investigation showed a significant linear trend component across days which would support the observation that pilots generally learn (i.e., show improvement) across the three days of testing (see Figure 7). This finding, however, must be interpreted in the presence of significant first-order interactions that are graphically presented in Figures 8 and 9 which show mean percent TOT by condition for X and Y respectively across days. These interactions, although significant, are not readily interpretable and probably reflect variability in the form of pilot difference interactions more than anything else. As a side note, the ANOVA profiles of the X and Y lateral measures were essentially identical. Similar results with measures that are highly correlated, such as the X and Y measures, indicate the presence of a good, stable measure. At the same time, TOT scores are sensitive to changes in pilot strategies. For example, when pilots try to fly the tasks within the prescribed tolerances, they will change their performance strategies based on the outcome of previous trials.

Table 12 also presents intercorrelations between X, Y and altitude TOT measures for the air taxi task with the four sickness indices. Obviously, poorer performance in the form of higher TOT scores are associated with higher ratings on the sickness indices.

TABLE 12. CORRELATIONS BETWEEN SICKNESS INDICES AND CATEGORY GROUPINGS, SIMULATOR PERFORMANCE DATA, AND TASK DIFFICULTY RATINGS

							AIR TAXI			SLALOM			
	Illness Composite-Subj	Illness Composite-E1	Illness Composite-E2	Illness Composite-$\bar{E}$	Intrahop Illness-Subj	MSQ Composite	Lateral (X) TOT	Lateral (Y) TOT	Altitude TOT	Lateral (Y) TOT	Airspeed TOT	Altitude TOT	Hit Rate (%)
Illness Composite-E1	.76												
Illness Composite-E2	.62	.87											
Illness Composite-$\bar{E}$	.71	.97	.97										
Intrahop Illness-Subj	.71	.59	.47	.55									
Illness Composite-Indp	.45	.64	.49	.59									
MSQ Composite	.58	.58	.55	.58	.47								
Post-hop Equil.- Pref	-.41	-.35	-.23	-.30	-.19	-.32							
Post-hop Equil.- NPref	-.40	-.36	-.29	-.34	-.20	-.25							
Flight Hours	.39	.33	.41	.41	.25	.29							
Simulator Hours	-.45	-.46	-.38	-.44	-.33	-.34							
AIR TAXI (AT)													
Lateral (X) TOT	.22	.26	.16	.22	.21	.33							
Lateral (Y) TOT	.22	.30	.12	.22	.17	.29							
Altitude TOT	.41	.30	.24	.28	.39	.23							
SLALOM (S)													
Lateral (Y) TOT	-.07	-.07	.02	-.03	.08	.03							
Airspeed TOT	.20	.15	.18	.17	.19	.29							
Altitude TOT	.00	-.04	-.03	-.04	.13	.19							
Hit Rate (%)	.16	.09	-.02	.04	.13	.01							
Task Difficulty- AT	.40	.38	.23	.32	.48	.34	.36	.41	.21				
Task Difficulty- S	.44	.40	.24	.33	.40	.50				.00	.23	.10	.01

$r = .195$, $p < .05$

$r = .274$, $p < .01$

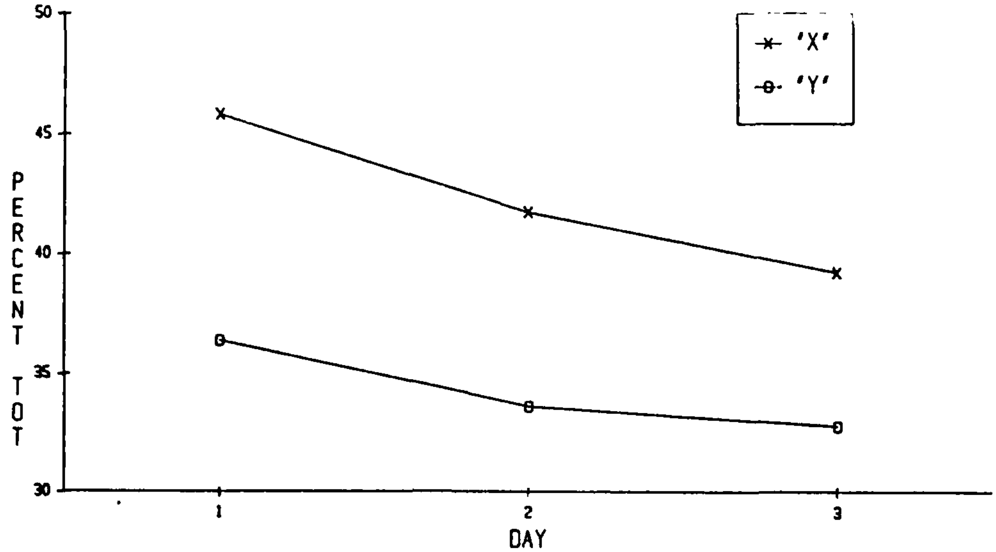


Figure 7. Mean Percent TOT for Two Lateral Performance Measures Across Days - Air Taxi

SLALOM. Tables 16-19 present summary ANOVA tables for lateral, altitude, and airspeed TOT measures as well as hit rate percentage, which gives an index of accuracy within the slalom. In this case, lag had a marginal effect on altitude ($F(2,40) = 3.01$, $p < .07$) and airspeed ($F(2,40) = 2.49$, $p < .10$). Pairwise mean comparisons for altitude performance revealed that the SH-3 condition (i.e., long lag, long variability) was worse than the CH-53E condition (i.e., moderate lag, moderate variability). In addition, there were no differences between SH-3 and VTRS (i.e., short lag, short variability) conditions or between CH-53E and VTRS conditions. Pairwise mean comparisons for airspeed performance revealed that the SH-3 condition was worse than either the CH-53E or VTRS condition. There was no difference between the CH-53B and VTRS lag conditions.

Lateral and airspeed TOT measures showed a significant day effect ($F(2,40) = 8.76$, $p < .001$; $F(2,40) = 4.79$, $p < .025$). As with the air taxi task, the significant linear trend component together with the observations taken from Figure 10 show that pilots' performance generally improved with time.

The significant day X lag interaction for the lateral TOT measure was present because, unlike pilots in the CH-53E and VTRS conditions which demonstrated increasingly better performance from Day 1 through to Day 3, pilots' performance in the SH-3 condition actually worsened from Day 1 to Day 3 (see Figure 11).

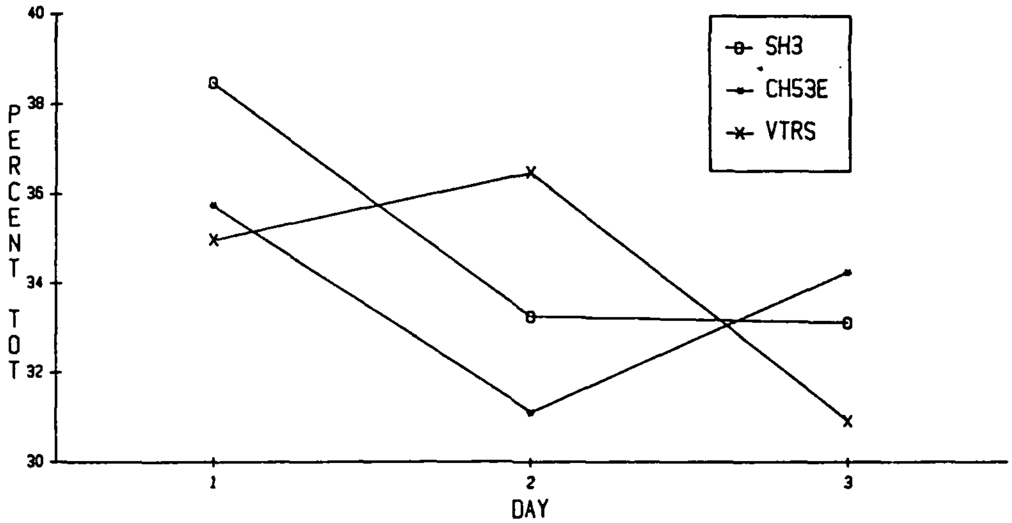


Figure 8. Mean Percent TOT for Lateral Performance by Condition Across Days - Air Taxi (X)

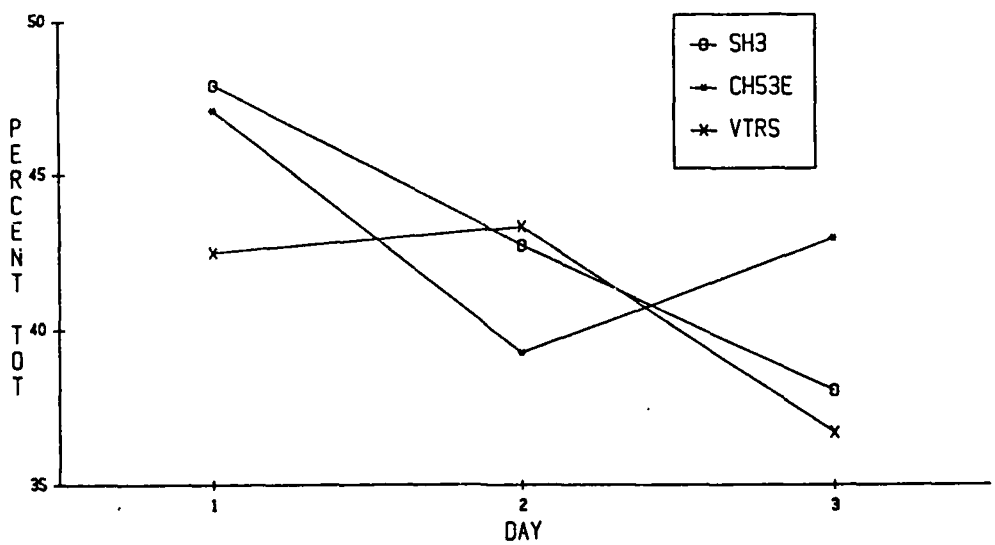


Figure 9. Mean Percent TOT for Lateral Performance by Condition Across Days - Air Taxi (Y)

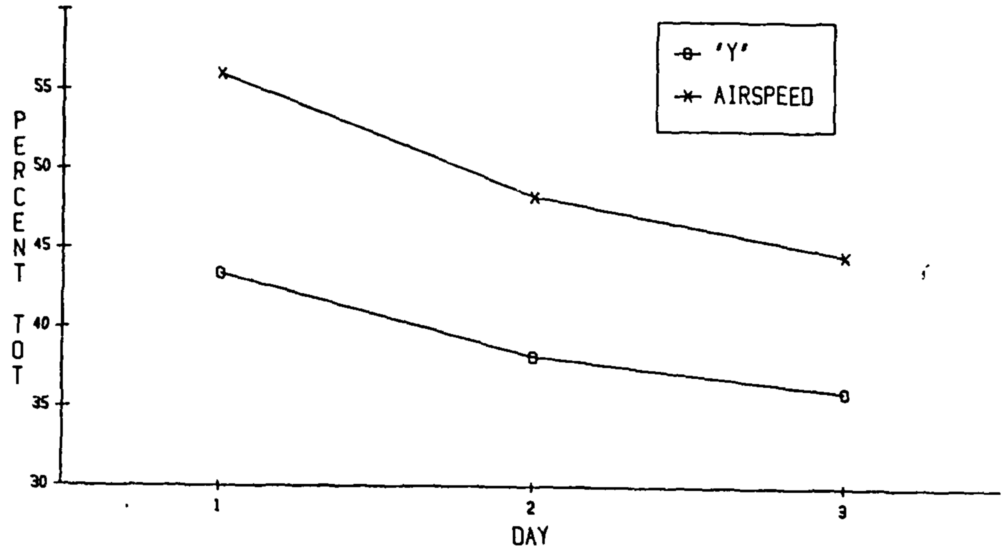


Figure 10. Mean Percent TOT for Lateral and Airspeed Performance Across Days - Slalom

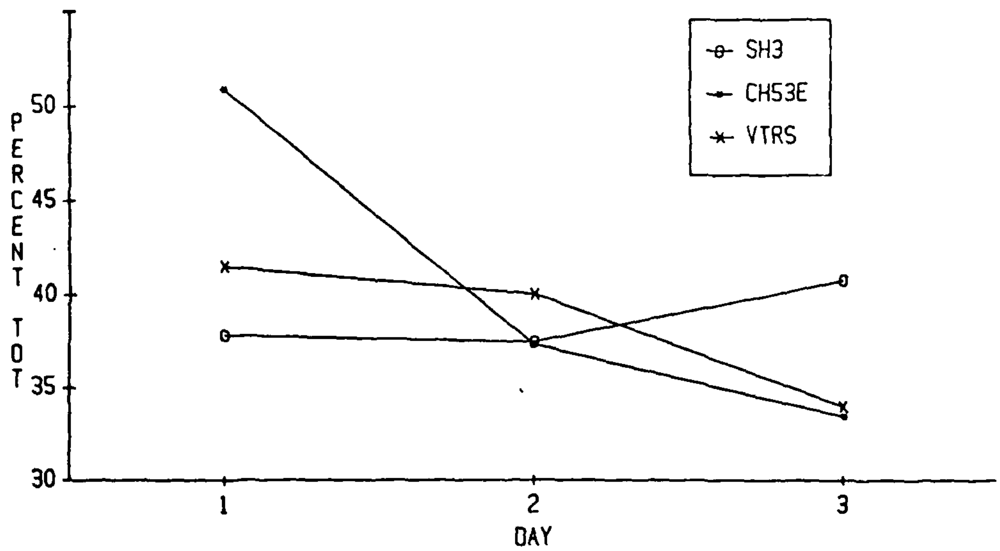


Figure 11. Mean Percent TOT for Lateral Performance by Condition Across Days - Slalom

Finally, unlike the finding for the air taxi task that poorer simulator performance is associated with higher sickness ratings, no such relationship appeared between performance on the slalom task and the same sickness indices.

ADDITIONAL ANALYSES

We were concerned that as a result of being ill on Day 1 the pilots would perform differently through the rest of the experiment than those pilots who were not ill enough to exit the simulator. Therefore, those pilots ($\underline{n}$=11) who were ill on Day 1 (i.e., they requested to leave the simulator before their hop was finished) were charted across the remaining two days of the experiment whether or not they were ill. These pilots were then compared with ones who were categorized as not ill ($\underline{n}$=13) based on the fact that they were able to complete their hop on all three days. It would, however, be misleading to state that all pilots who finished their first hop were unaffected by illness symptoms.

The subsequent analysis compared early exit vs no exit groups on flight variables for both tasks. On all measures (lateral and altitude TOT scores) for the air taxi task, the exit group performed significantly <u>better</u>. For the slalom task, airspeed performance was better for the exit group; however, those labeled as no exit performed better on the lateral and hit rate measures (see Figures 12 and 13).

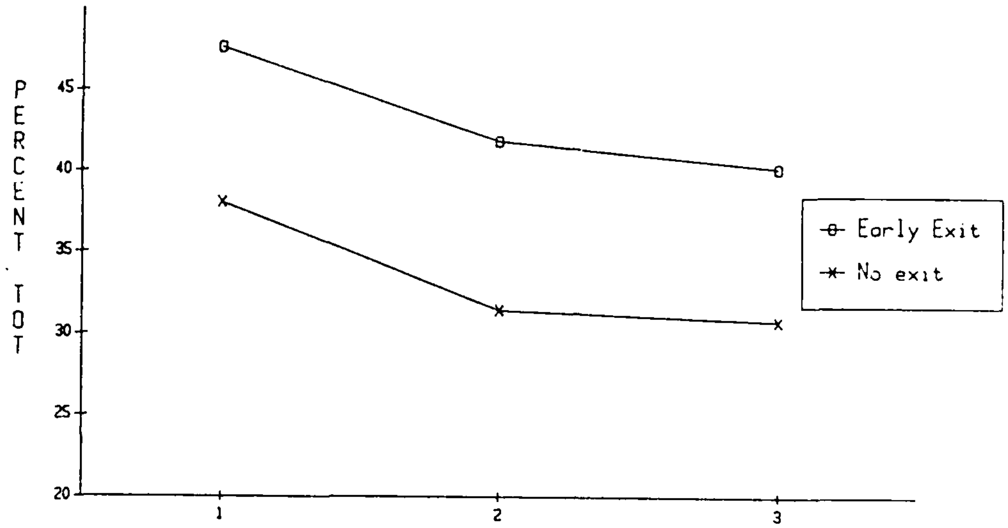


Figure 12. Mean Percent TOT for Lateral Performance for Exit/No Exit Groups - Slalom

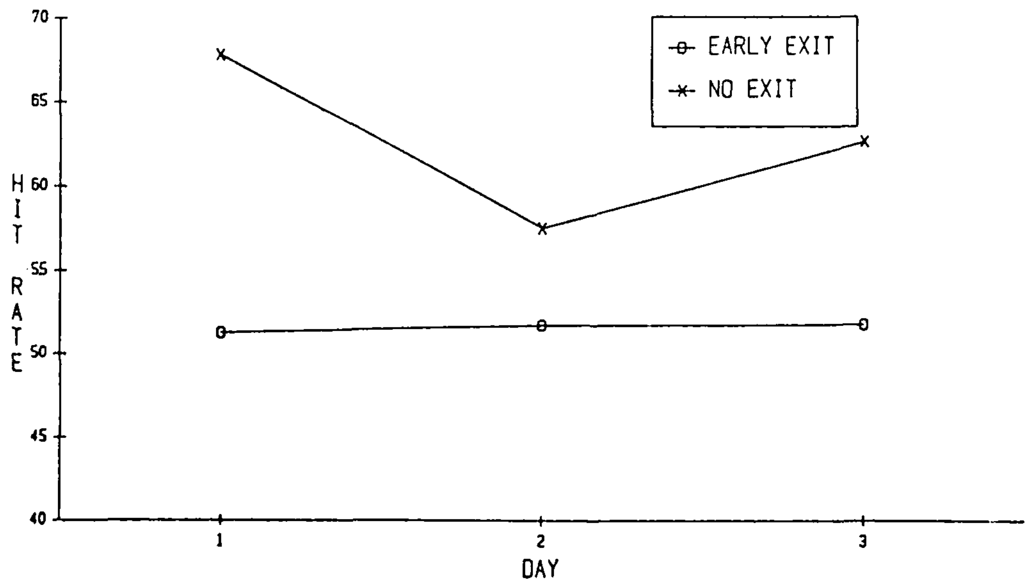


Figure 13. Hit Rate Percentage for Exit/No Exit - Slalom

TABLE 13. ANALYSIS OF VARIANCE FOR TOT LATERAL PERFORMANCE (X) - AIR TAXI TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	508.21	2	254.11	9.75	.001
Lag (L)	66.37	2	33.19	1.27	
D x L	313.69	4	78.42	3.01	.05
Pilots	2688.75	23	116.90	4.49	.001
Residual	1041.89	40	26.05		
TOTAL	4618.91	71			

TABLE 14. ANALYSIS OF VARIANCE FOR TOT LATERAL PERFORMANCE (Y) - AIR TAXI TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	219.90	2	109.95	8.54	.001
Lag (L)	25.35	2	12.68	0.98	
D x L	192.03	4	48.01	3.73	.05
Pilots	1135.03	23	49.35	3.38	.001
Residual	515.01	40	12.88		
TOTAL	2087.32	71			

TABLE 15. ANALYSIS OF VARIANCE FOR TOT
ALTITUDE - AIR TAXI TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	4865.29	2	2432.65	11.41	.001
Lag (L)	7.18	2	3.59	0.02	
D x L	265.17	4	66.29	0.31	
Pilots	5584.08	23	242.79	1.14	
Residual	8529.52	40	213.24		
TOTAL	19251.24	71			

TABLE 16. ANALYSIS OF VARIANCE FOR TOT
LATERAL PERFORMANCE (RUNWAY) - SLALOM TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	655.00	2	327.50	8.76	.001
Lag (L)	61.19	2	30.60	0.82	
D x L	966.25	4	241.56	6.46	.001
Pilots	9960.82	23	433.08	11.59	.001
Residual	1494.59	40	37.36		
TOTAL	13137.85	71			

TABLE 17. ANALYSIS OF VARIANCE FOR TOT
ALTITUDE - SLALOM TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	612.12	2	306.06	2.96	
Lag (L)	623.45	2	311.73	3.01	.07
D x L	1070.51	4	267.63	2.59	
Pilots	8245.76	23	358.51	3.46	.001
Residual	4140.81	40	103.52		
TOTAL	14692.65	71			

TABLE 18. ANALYSIS OF VARIANCE FOR TOT
AIRSPEED - SLALOM TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	1581.71	2	780.86	4.79	.025
Lag (L)	822.33	2	411.17	2.49	.10
D x L	927.56	4	231.89	1.40	
Pilots	11936.59	23	518.98	3.14	.001
Residual	6605.65	40	165.14		
TOTAL	21873.84	71			

TABLE 19. ANALYSIS OF VARIANCE FOR
HIT RATE (%) - SLALOM TASK

Source	SS	df	MS	F	<p
Within Subjects					
Days (D)	192.35	2	96.18	0.50	
Lag (L)	338.71	2	169.36	0.88	
D x L	1570.10	4	392.53	2.03	
Pilots	11379.10	23	494.74	2.56	.01
Residual	7735.79	40	193.39		
TOTAL	21216.05	71			

SECTION V

DISCUSSION

Visual lag had no effect on illness in this experiment. Although the longer lags were somewhat disruptive of performance, there was no evidence that they contributed to illness. This result was somewhat surprising in that the relatively large asynchronous lags present in the 2F64C and 2F120 simulators were considered prime suspects in the high illness rates associated with them. The range of lag and asychrony studied was fairly large (126 ± 17 msec to 215 ± 70 msec), and this would suggest that within the range of operationally useful simulators, visual asychrony does not appreciably contribute to simulator sickness. Lags approaching 300 msec in flight simulators become untenable from an operational or acceptability standpoint as they cause the simulation to be too unrealistic and/or too difficult to fly.

However, caution must be exercised in generalizing these results. First, the experiment examined only two tasks, and while these were somewhat generic in nature, they were also utilized because of their nauseogenic properties. It is possible that these results would be different with other types of tasks. Second, the experiment was performed without an inertial motion platform. Thus, no information was obtained regarding lag or asynchrony between visual and inertial systems. On the other hand, the experiment clearly demonstrated that simulator sickness can be generated by the visual display alone, and thus some elements or factors in the display are (potentially) responsible for high incidences of simulator sickness. These elements probably do not include visual lag asynchrony (within the limits studied) but may include certain inherent task characteristics.

Almost all of the pilots in this study experienced symptoms traditionally associated with motion sickness -- generally irrespective of the three visual lags utilized to study the phenomenon. Moreover, 46% of the pilots felt so ill on the first day of the experiment that they requested to leave the simulator prior to the end of their hop. However, the pilots, if adversely affected on the first hop, showed an ability to rapidly adapt to the ill effects of the simulation on following hops (days). Unfortunately, the Motion History Questionnaire (MHQ) as it currently exists was not successful in identifying individual susceptibility to illness. Coefficients between the MHQ and the paper-and-pencil sickness indices ranged from 0.03 to -0.28 -- values of little practical significance. There is currently an effort underway which will further investigate the psychometric properties of the form -- modifications are expected (Kennedy, Lenel, & Berbaum, in preparation). It may be fruitful to reevaluate the updated form using the data from this experiment.

Simulator performance on the air taxi task, while not affected by the lag, did show significant discrepancies between performance on the two measures of lateral performance (see Figure 7 and Section III, Simulator Performance Measurement). Further inspection revealed this difference to be due to the components of the visual scene. Specifically, while flying along the second and fourth legs (Y) on the helipad, pilots had more salient positioning information than when flying along the first and third legs (X). This

information was related to perspective transformations in simulated objects, including towers, buildings, other helicopters, and mountain sides (refer to Figure 1). It may be useful to investigate these differences further as they may relate to simulator sickness.

Performance on the slalom task, however, was affected in the predicted direction by the lag conditions. For altitude and airspeed control, performance was poorest for the lag condition with the longest lag duration and greatest lag variability (i.e., SH-3 condition). For both tasks, pilots generally improved their ability to fly within the prescribed tolerances and reported the tasks to become easier with practice. There was large variability in the time since each pilot's last simulator hop (i.e., 2 weeks-4 years). The mean time since the last simulator hop was approximately 10 months, indicating that these pilots recently spent a large portion of their time in their respective aircraft. This may explain poorer simulator flight performance on Day 1, but the question still remains if this was a contributing factor to the large incidence of illness on Day 1. Based on the predictions generated by the cue conflict theory, the latter is a reasonable conclusion. Also, the pilots flew aircraft ranging from utility UH-60s and UH-1s to ASW SH-3s and SH-60Bs. Although large within-subjects variability could not be statistically attributed to specific sources, it is likely that aircraft (simulator) unfamiliarity contributed to this source of variance. If pilots do indeed quickly learn the nuances of unfamiliar simulators as Westra and Lintern (1985) found, then performance should improve with practice. The data for this experiment generally supports this contention. In addition, learning was paired with adaptation to physical discomforts produced by the simulator. This finding, in the form of significant correlation coefficients, was supported only for the air taxi task. These results may indicate that even though pilots learned the requirements of the slalom task, its intrinsic nauseogenic nature made adaptation slower and inconsistent across pilots.

The results suggested that greater flight experience and less simulator experience lead to greater illness. This supports predictions that would be made from the cue conflict theory of simulator sickness (Kennedy, Berbaum, & Frank, 1984). This theory states that motion information from the visual and vestibular systems may be in conflict with expected values based on a neural store. These expected values reflect past experience(s). Any symptoms of illness, however, quickly dissipate with time without regard to the amount of flight or simulator experience.

An important contribution of the present study is the demonstration that the air taxi and slalom tasks were moderately difficult flight tasks that proved successful for inducing and studying simulator sickness. Also, the illness and MSQ scales were easy to administer instruments that proved sensitive to measuring changes in illness (i.e., adaptation). The strong experimenter/pilot and interrater agreement in rating illness demonstrates that there is consistency when observing and documenting the simulator sickness phenomenon. Finally, the relationship between the sickness ratings and postural equilibrium performance lends support to the notion that simulator exposure may result in aftereffects that are not unlike the aftereffects of other motion devices, but their relevance to safety is not fully understood and requires cautious interpretation.

The photographic rating procedure was successful to an extent, but was not without shortcomings. Pilots, onsite raters, and naive raters reported statistically similar composite illness ratings. Basically, this answers the question of whether any symptom(s) that lead the pilot and raters to similar ratings are overt enough to be captured on film. But, it is unclear whether particular symptoms received different weighting by individual raters. For example, even though the pallor and sweating dimensions were defined, the raters could have made an overall illness rating based on any criteria. Although this is an interesting application, we feel that the photographs do not substantially add useful information to the onsite rating process. In addition, the interrater reliability index was markedly less than the 0.87 value obtained by using the onsite raters, suggesting that the photograph rating procedure injected a somewhat larger degree of unwanted random variability into the assessment process. This finding also reinforces the notion that the onsite raters probably used information besides facial pallor and sweating to arrive at a rating (i.e., posture, tilt, walking behavior, etc.).

In any case, some important results must be identified. First, a measurable degree of objectivity can be inserted into an otherwise subjective effort. Secondly, even though experimentally naive raters are preferred in such a rating process, it is not often feasible. The results of the photographic assessment demonstrate that raters can be quickly and easily trained to recognize symptoms associated with simulator (motion) sickness. However, onsite raters may make more reliable ratings because they take a holistic (i.e., "Does he look sick?") approach to the rating process.

As previously discussed, two additional factors were added to the experimental design at the midpoint of the experiment. First, the exposure time was increased by 35% (from 40 to 60 minutes) in an effort to realize more incidence of illness. Second, there was concern that the air taxi task, by itself, was an easy task. Therefore, a small amount of random turbulence and 5 knots of wind were inserted into the task translating into tailwind, crosswind, and headwind components, respectively, as pilots performed the task. The first factor apparently did not result in incremental incidence of illness. Those pilots who were predisposed to illness continued to exit the simulator within the first 40 minutes. Moreover, the fact that the intrahop illness ratings were significantly correlated with end-of-hop illness ratings ($r=.71$, $p < .01$) seems to indicate that the second 20 or 30 minutes of the task were not adding much to the overall degree of illness. There was, however, a higher incidence of illness among the pilots who performed with the added turbulence and wind, suggesting that greater pilot workload could increase the nauseogenic content of the task.

SECTION VI

CONCLUSION AND RECOMMENDATIONS

The initial incidence of sickness among pilots used for this experiment was almost 50% using the criterion of early exit/no exit on Day 1. Overall, the magnitude of illness as measured by the paper-and-pencil instruments decreased with each subsequent day. Pilots may have been operationally conditioned to avoid the outside visual scene as much as possible. This strategy, combined with practice, is probably responsible for the observed adaptation. Also, for this study, the initial 40 minutes of exposure was the most provocative, and adding an additional 20 minutes of exposure failed to provide greater incidence of illness. Finally, lag had no effect on any of the sickness ratings.

The effect of asynchronous lag on in-simulator flight task measures, while not consistent, was typically associated with poorer task performance. Pilots were almost unanimously aware of the two longest lags (CH-53E and SH-3) and reported that the jerky visual quality of the scene was bothersome -- especially in the precision air taxi task. However, simulator performance did not uniformly suffer. Even so, because of the noticeable and generally disruptive nature of the longer lag conditions, it is recommended that they be avoided in operational simulators. Although they may not contribute to illness <u>per se</u>, they could interact with other factors to increase illness. In any case, the longer lags appear to be excessive from both a performance and pilot opinion standpoint. Further study is necessary to establish reasonable delay magnitudes with respect to specific task and transfer-of-training issues.

Based on the results presented in this study, several observations and recommendations are offered:

- In this study, adaptation was a strong effect. It could be argued that, if left alone, individuals are flexible enough to "fix themselves"; however, if the definition of simulator sickness is synonomous with a bad simulator design, then adaptation should not be the solution.

- The air taxi and slalom tasks appeared to both be successful conduits for studying simulator sickness. Tasks chosen for such a purpose must attempt to concentrate the pilot's focus mainly on the outside visual field. It may be necessary to cover or obstruct instruments in order to force the pilot's gaze mainly out of the cockpit.

- Hops of 60 minutes in length do not increase the incidence of sickness compared to hops of 40 minutes.

- Sickness can be easily and reliably measured by a simple to administer and noninvasive rating instrument such as the Global Illness Rating Scale. Moreover, using pilot's self-reported assessments of illness as a criterion, assessment by onsite raters are highly predictive.

- Postural difficulties are directly related to the severity of illness. In addition, those pilots who requested to leave the simulator prematurely also completed the posthop section of the MSQ at half-hour intervals up to two hours. The symptoms reported by the pilots in this group disappeared after a one-half hour posthop. Based on this finding, pilots should be detained at least that long before being dismissed. Common sense tells us that care should still be taken even when the severity of illness is essentially nil.

- The data generated from this experiment does fit within the cue conflict theory of simulator sickness. Therefore, it is recommended that pilots with a relatively high number of flight hours (e.g., > 1000) be afforded more initial familiarization time in the simulator followed by a short rest period before any lengthy hop is begun. Overall, this procedure should be welcomed by all pilots. Kennedy, Lilienthal, Berbaum, and Dunlap (1986) provide additional guidelines that should be followed to lessen either the severity and/or incidence of illness. These guidelines focus on areas such as the duration and nature of exposure and procedural characteristics of the training syllabus.

- Finally, one of the most frequent comments made by the pilots was an "uneasiness" caused by a lack of motion cues. We suggest that future research should focus on the ability of a motion base and/or motion seat to mitigate these effects.

REFERENCES

Alexander, S. J., Cotzin, M., Hill, C. J., Jr., Ricciuti, E. A., & Wendt, G. R. Wesleyan University studies of motion sickness: I. The effects of variation of time intervals between accelerations upon sickness rates. Journal of Psychology, 1945, 19, 49-62. (a)

Alexander, S. J., Cotzin, M., Hill, C. J., Jr., Ricciuti, E. A., & Wendt, G. R. Wesleyan University studies of motion sickness: III. The effects of various accelerations upon sickness rates. Journal of Psychology, 1945, 20, 3-8. (b)

Alexander, S. J., Cotzin, M., Hill, C. J., Jr., Ricciuti, E. A., & Wendt, G. R. Wesleyan University studies of motion sickness: VI. Prediction of sickness on a vertical accelerator by means of a motion sickness history questionnaire. Journal of Psychology, 1945, 20, 25-30. (c)

Alexander, S. J., Cotzin, M., Hill, C. J., Jr., Ricciuti, E. A., & Wendt, G. R. Wesleyan University studies of motion sickness: VII. The effects of sickness upon performance. Journal of Psychology, 1945, 20, 31-39. (d)

Baddeley, A. D. A three-minute reasoning test based on grammatical transformation. Psychonomic Science, 1968, 10, 341-342.

Baron, S.. An optimal control model analysis of data from a simulated hover task. Paper presented at the 18th Annual Conference on Manual Control, Dayton, OH, June 1982.

Berbaum, K. S., Kennedy, R. S., Dunlap, W. Guidelines for reduction in symptoms of simulator sickness in simulators at MCAS New River. Unpublished manuscript. Orlando, FL: Essex Corporation, 1986.

Bittner, A. C., Jr., Carter, R. C., Kennedy, R. S., Harbeson, M. M., & Krause, M. Performance Evaluation Tests for Environmental Research (PETER): Evaluation of 112 measures. Perceptual and Motor Skills, 1986, 63, 683-708.

Bittner, A. C., Jr., Smith, M. G., Kennedy, R. S., Staley, C. F., & Harbeson, M. M. Automated Portable Test (APT) System: Overview and prospects. Behavior Research Methods Instruments & Computers, 1985, 17(2), 217-221.

Bray, D. W., & Grant, D. L. The assessment center in the measurement of potential for business management. Psychological Monographs, 1966, 80(17) (Whole No. 625), 1-27.

Browder, B., & Butrimas, S. Memorandum: Test and evaluation of device 2F64C motion system. Orlando, FL: Naval Training Equipment Center, Code N-732, 1983.

Browder, B., & Butrimas, S. Memorandum: Throughput testing of device 2F64C (SH-3) helicopter) visual system. Orlando, FL: Naval Training Equipment Center, Code N-732, 1984.

Caro, P. W. Some factors influencing Air Force simulator training effectiveness. Alexandria, VA: Human Resources Research Organization, HumRRO-TR-77-2, March 1977.

Carter, R. C., Kennedy, R. S., & Bittner, A. C., Jr. Grammatical reasoning: A stable performance yardstick. Human Factors, 1981, 23, 587-591.

Colehour, J. K., & Graybiel, A. Biochemical changes occurring with adaptation to accelerative forces during rotation (Joint Report No. NAMI-959). Pensacola, FL: NASA/U.S. Naval Aerospace Medical Institute, April 1966.

Coussens, W. Study: Evaluation of effects of work at altitude and mountain sickness on performance. R. S. Kennedy (Ed.). In APTS Newsletter, 1(1), Orlando, FL: Essex Corporation, 1985.

Crosby, T. N., & Kennedy, R. S. Postural disequilibrium and simulator sickness following flights in a P-3C operational flight trainer. Preprints of the 53rd Annual Scientific Meeting of the Aerospace Medical Association, Bal Harbor, FL: 10-13 May 1982.

Department of Defense. Military Standard 1558. Six-degree-of-freedom motion system requirements for air crewmember training simulations (MILSTD-1558). Washington, DC, 1974.

Dolezal, H. Living in a world transformed. New York: Academic Press, 1982.

Evans, R. M., Scott, P. G., & Pfeiffer, M. G. SH-3 helicopter flight training: An evaluation of visual and motion simulation in device 2F64C (Technical Report No. 161). Orlando, FL: Naval Training Equipment Center (Code 1), 1984.

Frank, L. H., Kennedy, R. S., McCauley, M. E., & Kellogg, R. S. Simulator sickness: Reaction to a transformed perceptual world. I. Scope of the problem. NAVTRAEQUIPCEN TN-65. Orlando, FL: Naval Training Equipment Center, 1983.

Fregly, A. R. Vestibular ataxia and its measurement in man. In H. H. Kornhuber (Ed.). Handbook of sensory physiology Vol. VI/2: Vestibular system; Part 2: Psychophysics, applied aspects and general interpretations. Berlin, GERMANY: Springer-Verlag, 1974.

Fregly, A. R., & Kennedy, R. S. Comparative effects of prolonged rotation at 10 RPM on postural equilibrium in vestibular normal and vestibular defective human subjects. Aerospace Medicine, 1965, 36(12), 1160-1167.

Fregly, A. R., Smith, M. J., & Graybiel, A. Revised normative standards of performance of man on a quantitative ataxia test battery. Acta Otolaryngolica, 1973, 75, 10-16.

Graybiel, A., & Fregly, A. R. A new quantitative ataxia test battery. Pensacola, FL: U.S. Naval School of Aviation Medicine, NASA Symposium on the Role of the Vestibular Organs in the Exploration of Space, 1965.

Graybiel, A., Wood, C. D., Miller, E. F., II, & Cramer, D. B. Diagnostic criteria for grading the severity of acute motion sickness. Aerospace Medicine, 1968, 39, 453-455.

Hardacre, L. E., & Kennedy, R. S. Some issues in the development of a motion sickness questionnaire for flight students. Aerospace Medicine, 1963, 34, 401-402.

Held, R., Efstathiou, A., & Greene, M. Adaptation to displaced and delayed visual feedback from the hand. Journal of Experimental Psychology, 1966, 72, 887-891.

Herndon, J. W. A visual technology research simulator for vertical takeoff and landing (VTOL) (NAVTRAEQUIPCEN IH-337). Naval Training Equipment Center, Orlando, FL, 1982.

Hill, J. C. A model of the human postural control system. Paper presented at the Joint National Conference on Major Systems, Anaheim, CA: 25-29 October 1971.

Hutchins, C. W., & Kennedy, R. S. Relationship between past history of motion sickness and attrition from flight training. Aerospace Medicine, 1965, 36, 984-987.

Kellogg, R. S., Castore, C., & Coward, R. E. Psychophysiological effects of training in a full vision simulator. Paper presented at the Human Factors Society Meeting, 1980.

Kennedy, R. S., Dutton, B., Lilienthal, M. G., Ricard, G. L., & Frank, L. H. Simulator sickness: Incidence of simulator aftereffects in 10 Navy flight trainers. Paper presented at the SAFE Symposium, Las Vegas, NV: 1984.

Kennedy, R. S., Dutton, B., Ricard, G. L., and Frank, L. H. Simulator sickness: A survey of flight simulators for the Navy. Long Beach, CA: Paper presented at the Aerospace Congress and Exposition, SAE, 1984.

Kennedy, R. S., Frank, L. H., McCauley, M. E., Bittner, A. C., Jr., Root, R. W., & Binks, T. A. Simulator sickness: Reaction to a transformed perceptual world. VI. Preliminary site surveys. Paper presented at the AGARD Conference No. 372, Williamsburg, VA, 1984.

Kennedy, R. S., & Graybiel, A. A comparison of the susceptibilities of three groups of Naval aviation personnel to symptomatology of motion sickness. Paper presented at the American Psychological Association Meetings, Philadelphia, PA: September 1963. (a)

Kennedy, R. S., & Graybiel, A. Pensacola diagnostic categorization worksheet for the evaluation of vestibular sickness. NAVSCOLAVNMED Form No. 6500/25, 1963. (b)

Kennedy, R. S., & Graybiel, A. The Dial Test: A standardized procedure for the experimental production of canal sickness symptomatology in a rotating environment. Joint Report No. NSAM-930. Pensacola, FL: NASA/US Naval School of Aviation Medicine, 1965.

Kennedy, R. S., Lenel, J., & Berbaum, K. S. Revised Motion History Questionnaire. In preparation.

Kennedy, R. S., Lilienthal, M. G., Berbaum, K. S., & Dunlap, W. P. Issues in simulator sickness. Paper presented at the International Conference: Advances in Flight Simulation Visual and Motion Systems. London, England, April 1986.

Kennedy, R. S., Merkle, P. J., & Lilienthal, M. G. A comparison of postural equilibrium effects following exposure to different ground-based flight trainers. Paper presented at the SAFE Symposium, Las Vegas, NV, 1985.

Kennedy, R. S., Wilkes, R. L., Lane, N. E., & Homick, J. L. Preliminary evaluation of a microbased repeated-measures testing system. Paper presented at the 56th Annual Meeting of the Aerospace Medical Association, San Antonio, TX, 1985.

Lentz, M., & Guedry, F. F. Motion sickness susceptibility: A comparison of laboratory tests. Aviation, Space and Environmental Medicine, 1978, 49, 1281-1288.

Lilienthal, M. G., & Merkle, P. J., Jr. Simulator sickness in flight simulators: A case study. Transportation Research Record 1059: Vehicular Simulation and User Behavioral Studies. Transportation Research Board, National Research Council, Washington, DC, 1986.

McCauley, M. E., (Ed.). Simulator sickness: Proceedings of a workshop. National Academy of Sciences/National Research Council/National Academy of Sciences, Washington, DC: Committee on Human Factors, 1984.

McCauley, M. E., & Kennedy, R. S. Recommended human exposure limits for very-low frequency vibration (PMTC-76-36). Point Mugu, CA: Pacific Missile Test Center, 1976.

McClure, J. A., & Fregly, A. R. Effect of environmental temperature on sweat onset during motion sickness. Aerospace Medicine, 1972, 43(9), 959-967.

McGuinness, J., Bouwman, J. H., & Forbes, J. M. Simulator sickness occurrences in the 2E6 Air Combat Maneuvering Simulator (ACMS). Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 80-C-0135-4500-1, 1981.

Miller, J. W., & Goodson, J. E. A note concerning "motion sickness" in the 2FH2 Hover Trainer. Research Project NM 170 111; Subtask 3, Report 1. Pensacola, FL: Naval School of Aviation Medicine, 1958.

Money, K. E. Motion sickness. Physiological Reviews, 1970, 50(1), 1-39.

Money, K. E. Flight simulator motion sickness in the Aurora CP140 FDS (Technical Communication No. 80-C-44). Downsview, Ontario, Canada: Defence and Civil Institute of Environmental Medicine, 1980.

Orlansky, J., & String, J. Cost effectiveness of flight simulators for military training. Volume I: Use and effectiveness of flight simulators, IDA Paper P-1275. Alexandria, VA: Institute for Defense Analyses Science and Technology Division, 1977. (a)

Orlansky, J., & String, J. Cost effectiveness of flight simulators for military training. Volume II: Estimating costs of training in simulators and aircraft, IDA Paper P-1275. Alexandria, VA: Institute for Defense Analyses, 1977. (b) (AD A052801)

Rapin, I., Costa, L. D., Mandel, I. J., & Fromowitz, A. J. Effect of varying delays in auditory feedback on key-tapping of children. Perceptual and Motor Skills, 1963, 16, 489-500.

Reason, J. T. Motion sickness and adaptation: A neural mismatch model. Journal of the Royal Society of Medicine, 1978, 71, 819-829.

Reschke, M. F., Homick, J. L., Ryan, P., & Moseley, E. C. Prediction of the space adaptation syndrome. Preprints of the AGARD Aerospace Medical Panel Symposium, Williamsburg, VA: 30 April-4 May 1984.

Ricard, G. L., Norman, D. A., & Collyer, S. C. Compensating for flight simulator CGI system delays. Paper presented at the 9th NTEC/Industry Conference, Orlando, FL, 1976.

Ricard, G. L., & Puig, J. A. Delay of visual feedback in aircraft simulators. Applied Psychology, 1977, 59, 2250-2258.

Ryan, L. E., Scott, P. G., & Browning, R. F. The effects of simulator landing practice and the contribution of motion simulation to P-3 pilot training (TAEG Report No. 63). Orlando, FL: Training Analysis and Evaluation Group, 1978.

Schiflett, S., Bowles, P., & Haswell, S. Effects of hypoxia on performance. Paper presented at the USAF School of Aerospace Medicine "Operation Problems in Aerospace Medicine," San Antonio, TX, 1985.

Smith, K. U. Special review: Sensory feedback analysis in medical research. I. Delayed sensory feedback in behavior and neural function. American Journal of Physical Medicine, 1963, 42(2), 49-84.

Steele, J. E. The symptomatology of motion sickness. Proceedings of the Fourth Symposium on the Role of the Vestibular Organs in Space Exploration (NASA SP-187). Naval Aerospace Medical Institute, 24-26 September 1968, 89-96.

Stockwell, C. W., Koozekanari, S. H., & Barin, K. A Physical model of human postural dynamics. Annals New York Academy of Sciences, 1981, 722-730.

Thomley, K. E., Kennedy, R. S., & Bittner, A. C., Jr. Development of postural equilibrium tests for examining simulator aftereffects. Preprints of the Annual Aerospace Medical Association Meeting, San Diego, CA: Manuscript submitted for publication, 1986.

Welch, R. B. Perceptual modification: Adapting to altered sensory environments. New York: Academic Press, 1978.

Westra, D. P., & Lintern, G. Simulator design features for helicopter landing on small ships. I. A performance study (NAVTREQUIPCEN 81-C-0105-13). Orlando, FL: Naval Training Equipment Center, September 1985.

Wiker, S. F., Kennedy, R. S., McCauley, M. E., & Pepper, R. L. Reliability, validity and application of an improved scale for assessment of motion sickness severity (CG-D-29-79). Washington, DC: Department of Transportation, 1979. (a)

Wiker, S. F., Kennedy, R. S., McCauley, M. E., & Pepper, R. L. Susceptibility to seasickness: Influence of hull design and steaming direction. Aviation, Space, and Enviromental Medicine, 1979, 50, 1046-1051. (b)

APPENDIX A

GLOBAL ILLNESS RATING SCALE

Name ______________________________

Date ______________________________

Instructions: Please rate yourself on your overall appearance. A mirror is available to help you make an accurate rating. Please circle the appropriate number.

PREHOP

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
appear                         appear                           appear
normal                    moderately ill                   extremely ill
```

POSTHOP

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
appear                         appear                           appear
normal                    moderately ill                   extremely ill
```

APPENDIX B

TASK DIFFICULTY RATING SCALE

S Name ____________________ Condition ____________________

Day (1, 2, 3) ______

Instructions. Please circle the number that most closely represents your judgments regarding the tasks.

1. How much attention and effort did the slalom task require:

1	2	3	4	5	6	7
Not very much			Moderate amount			Extreme amount

2. How much attention and effort did the air taxi task require:

1	2	3	4	5	6	7
Not very much			Moderate amount			Extreme amount

APPENDIX C

DESCRIPTION OF PHOTOGRAPH ASSESSMENT PROCESS

METHOD

Subjects. Six undergraduate college students (three males and three females) served as subjects. All subjects were paid $10 for their participation. The rating session lasted about 2-1/2 hours.

PROCEDURE

Training Session. Subjects were first given background information regarding the experiment proper. They were then instructed on the group discussion process. Next, the properties and use of the global illness rating scale presented in Appendix A was explained to the subjects who used an identical scale to provide ratings. Finally, the subjects were exposed to rating errors common to this type of exercise (e.g., central tendency, contrast). The actual training session consisted of five sets of pre post photographs extracted from the available pool. Of these, three sets showed obvious signs of illness, while the pre/post photographs of the remaining two sets were essentially identical.

Group Discussion. Each set of photographs were independently rated by each subject before they met as a group to review the ratings. Each subject read his/her overall rating for each pilot to an experimenter keeping a master list. If discrepancies of greater than two rating points were discovered, these subjects were required to discuss their reasons for the rating they gave. The subjects were also allowed to view the photographs again and direct attention to aspects of the photograph which led them to their initial rating. Space was provided on the rating form for additional comments or observation that the rater used in arriving at an initial rating. These comments were used as a basis for the discussion. Based on this discussion, subjects could, but were not required to, modify their ratings. This rationale is based on a variety of research (e.g., Bray & Grant, 1966) which suggest that raters will be in close agreement on their initial ratings and group discussion will bring them even closer.

Rating Session. Subjects rated 30 sets of photographs. If necessary, the ratings were discussed within the group discussion procedure described previously. Subjects were given 30 sets of 35mm color photographs of Navy and Marine pilots to examine. The sets of photographs were randomly chosen from the available pool of 72 sets. Photos with processing imperfections due to lighting, focus, or flash were eliminated from the initial pool.

APPENDIX D

MOTION SICKNESS QUESTIONNAIRE

Serial No.________

SIMULATOR SICKNESS SURVEY

This is a survey of simulator aftereffects being conducted for the Naval Training Equipment Center by Essex Corporation. The purpose of the survey is to determine the incidence of simulator aftereffects such as nausea or imbalance occurring in various types of flight simulators.

We appreciate your cooperation in obtaining information about this problem. The results of the study will be used to improve the characteristics of future simulators. Your responses will be held in confidence and used statistically. Although we ask for your name on this page, no information will be reported by name. This cover page will be removed and all data will be identified by a coded serial number.

Name____ Rank____________________

Date____ Squadron________________

Simulator_______________________________________

October 1985 Revision (VTRS)

Serial No.________

PREHOP BACKGROUND INFORMATION

Instructions: Please fill this page out BEFORE you go into the simulator. Fill in the blanks or circle the appropriate item.

1. Start time for your hop: ___________ Expected length of hop: ________

2. Your total amount of flight hours. _______________

3. Your total amount of simulator hours. _______________

4. How long has it been since your last flight IN THE AIRCRAFT?

____

5. How long has it been since your last flight IN A SIMULATOR? WHICH SIMULATOR?

____

GO TO THE NEXT PAGE AND CONTINUE

Serial No.________

PREHOP PHYSIOLOGICAL STATUS INFORMATION

Instructions: Please fill this out BEFORE you go into the simulator.

1. Are you in your usual state of fitness? YES NO

   If not, what is the nature of your illness (flu, cold, etc.)

   ____

2. Please indicate all medication you have used in the past 24 hours:

   a) NONE ____________________

   b) Sedatives or tranquilizers ____________________

   c) Aspirin, Tylenol, other analgesics ____________________

   d) Anti-histamines ____________________

   e) Decongestants ____________________

   f) Other (specify):____________________________ ____________________

3) How many hours sleep did you get last night? __________ (Hours)

   Was this amount sufficient? YES NO

GO ON TO THE NEXT PAGE

Serial No.________

PREHOP SYMPTOM CHECKLIST

Instructions: Please fill this out BEFORE you go into the simulator. Circle below if any sumptoms apply to you right now. (After your simulator exposure, you will be asked these questions again.)

1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Boredom	None	Slight	Moderate	Severe
4.	Drowsiness	None	Slight	Moderate	Severe
5.	Headache	None	Slight	Moderate	Severe
6.	Eye strain	None	Slight	Moderate	Severe
7.	Difficulty focusing	None	Slight	Moderate	Severe
8.	a. Salivation increased	None	Slight	Moderate	Severe
	b. Salivation decreased	None	Slight	Moderate	Severe
9.	Sweating	None	Slight	Moderate	Severe
10.	Nausea	None	Slight	Moderate	Severe
11.	Difficulty concentrating	None	Slight	Moderate	Severe
12.	Mental depression	No	Yes		
13.	"Fullness of the Head"	No	Yes		
14.	Blurred vision	No	Yes		
15.	a. Dizziness with eyes open	No	Yes		
	b. Dizziness with eyes closed	No	Yes		
16.	Vertigo	No	Yes		
17.	*Visual flashbacks	No	Yes		
18.	Faintness	No	Yes		
19.	Aware of breathing	No	Yes		
20.	**Stomach awareness	No	Yes		
21.	Loss of appetite	No	Yes		
22.	Increased appetite	No	Yes		
23.	Desire to move bowels	No	Yes		
24.	Confusion	No	Yes		
25.	Burping	No	Yes	No. of times	
26.	Vomiting	No	Yes	No. of times	
27.	Other				

* Visual illusion of movement or false sensations similar to aircraft dynamics, when not in the simulator or the aircraft.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

STOP HERE! The test director will tell you when to continue.

POSTHOP DATA

Serial No.________

POSTHOP SYMPTOM CHECKLIST

Instructions: Please fill this out AFTER you go into the simulator. Circle below if any symptoms apply to you right now.

1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Boredom	None	Slight	Moderate	Severe
4.	Drowsiness	None	Slight	Moderate	Severe
5.	Headache	None	Slight	Moderate	Severe
6.	Eye strain	None	Slight	Moderate	Severe
7.	Difficulty focusing	None	Slight	Moderate	Severe
8.	a. Salivation increased	None	Slight	Moderate	Severe
	b. Salivation decreased	None	Slight	Moderate	Severe
9.	Sweating	None	Slight	Moderate	Severe
10.	Nausea	None	Slight	Moderate	Severe
11.	Difficulty concentrating	None	Slight	Moderate	Severe
12.	Mental depression	No	Yes		
13.	"Fullness of the Head"	No	Yes		
14.	Blurred vision	No	Yes		
15.	a. Dizziness with eyes open	No	Yes		
	b. Dizziness with eyes closed	No	Yes		
16.	Vertigo	No	Yes		
17.	*Visual flashbacks	No	Yes		
18.	Faintness	No	Yes		
19.	Aware of breathing	No	Yes		
20.	**Stomach awareness	No	Yes		
21.	Loss of appetite	No	Yes		
22.	Increased appetite	No	Yes		
23.	Desire to move bowels	No	Yes		
24.	Confusion	No	Yes		
25.	Burping	No	Yes	No. of times	
26.	Vomiting	No	Yes	No. of times	
27.	Other				

* Visual illusion of movement or false sensations similar to aircraft dynamics, when not in the simulator or the aircraft.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Serial No.________

POSTHOP INFORMATION

Instructions: Please fill out this page AFTER you have completed your hop. If you are not sure about a certain answer please give your best estimate.

1. Percentage of time looking out the window for:

   A. Air taxi __________

   B. Slalom __________

2. Percentage of time NOT looking out the window for:

   A. Air taxi __________

   B. Slalom __________

3. Number of times the simulator was put on freeze: __________

4. Number of crashes: __________

5. Time spent in simulator: __________ The time now: __________

6. Did you have to wait long periods while in the simulator for any reason?

   Yes _____ No _____ HOW LONG? _____

7. Scene Disturbances:

   Describe any descriptive visual problems that you observed:

   ____

   ____

8. Describe any bothersome visual traits that you would like to see corrected.

   ____

   ____

APPENDIX E

MOTION HISTORY QUESTIONNAIRE

Serial No.________

SIMULATOR SICKNESS SURVEY

This is a survey of simulator aftereffects being conducted for the Naval Training Equipment Center by Essex Corporation. The purpose of the survey is to determine the incidence of simulator aftereffects such as nausea or imbalance occurring in various stages of flight simulators.

We appreciate your cooperation in obtaining information about this problem. The results of the study will be used to improve the characteristics of future simulators. Your responses will be held in confidence and used statistically. Although we ask for your name on this page, no information will be reported by name. This cover page will be removed and all data will be identified by a coded serial number.

Name ________________________ Rank ________________

Date ________________________ Squadron ________________

Simulator ________________________

10/85

ESSEX

MOTION HISTORY QUESTIONNAIRE

Serial # ___________

ESSEX

Motion History Questionnaire*

1. Approximately, how many total flight hours do you have?

Fixed Wing	Pilot/Co-Pilot	___________ Hours
	Crew	___________ Hours
Rotary Wing	Pilot/Co-Pilot	___________ Hours
	Crew	___________ Hours

2. How often would you say you get air sick?

   Always ___ Frequently ___ Sometimes ___ Rarely ___ Never ___

3. a) How many total flight simulator hours? _______________ Hours

   b) How many times have you been in this simulator? __________ Times

4. How much experience have you had at sea aboard ships or boats?

   Much ___ Some ___ Very Little ___ None ___

5. From your experience at sea, how often would you say you get seasick?

   Always ___ Frequently ___ Sometimes ___ Rarely ___ Never ___

6. Have you ever been motion sick under any conditions other than the ones listed so far?

   No ___ Yes ___. If so, under what conditions? _______________

   ____

7. In general, how susceptible to motion sickness are you?

   Extremely ___ Very ___ Moderately ___ Minimally ___ Not at all ___

*This questionnaire was prepared for research purposes only by R. S. Kennedy and M. E. McCauley. Scoring procedures are available from the authors at: Essex Corporation, 1040 Woodcock Rd., Suite 227, Orlando, FL 32803.

Serial # ________

ESSEX

8. Have you been nauseated FOR ANY REASON during the past 8 weeks?

No ___ Yes ___ If yes, explain ______________________________

____

9. When you were nauseated for any reason (including flu, alcohol, etc.), did you vomit:

Easily ___ Only with difficulty ___ Retch and finally vomited with great difficulty ___

10. If you vomited while experiencing motion sickness, did you:

a) Feel better and remain so? ___
b) Feel better temporarily, then vomit again? ___
c) Feel no better, but not vomit again? ___

d) Other - specify ______________________________

____

11. If you were in an experiment where 50% of the subjects get sick, what do you think your chances of getting sick would be?

Almost certainly would ___ Probably would ___ Probably would not ___ Almost certainly would not ___

12. Would you volunteer for an experiment where you knew that: (Please answer all three)

a) 50% of the subjects did get motion sick? Yes ___ No ___
b) 75% of the subjects did get motion sick? Yes ___ No ___
c) 85% of the subjects did get motion sick? Yes ___ No ___

13. Most people experience slight dizziness (not a result of motion) 3 to 5 times a year. The past year you have been dizzy:

more than this ___ the same as ___ less than ___ never dizzy ___

14. Have you ever had an ear illness or injury which was accompanied by dizziness and/or nausea? Yes ___ No ___

ESSEX

Serial # ___________

15. Listed below are a number of situations in which some people have reported motion sickness symptoms. In the space provided, check (a) your PREFERENCE for each activity (that is, how much you like to engage in that activity), and (b) any SYMPTOM(s) you may have experienced at any time, past or present.

SITUATIONS	PREFERENCE			SYMPTOMS													
	Like	Neutral	Dislike	Vomited	Nausea	Stomach awareness*	Increased salivation	Dizziness	Drowsiness	Sweating	Pallor	Vertigo	Awareness of breathing	Headache	Other symptoms	None
Aircraft																
Flight Simulator																
Roller Coaster																
Merry-Go-Round																
Other Carnival Devices																
Automobiles																
Long Train or Bus Trips																
Swings																
Hammocks																
Gymnastic Apparatus																
Roller/Ice Skating																
Elevators																
Cinerama or Wide-Screen Movies																
Motorcycles																

*Stomach awareness refers to a feeling of discomfort that is preliminary to nausea.

Serial # ___________

ESSEX

16. If you have ever experienced simulator sickness or discomfort (or any other aftereffect): what simulator was it?, what were the symptoms?, how long did they last?, and what do you think caused the problem?

____

____

____

____

____

END OF MOTION HISTORY QUESTIONNAIRE.